Journal of Medical Internet Research

Contents

XSL·FO

RenderX

XSL•FO
RenderX

Review

Short Papers

<u>Viewpoint</u>

The Inclusion of African-American Study Participants in Web-Based Research Studies: Viewpoint

Bekeela Watson[1*], MPH; Dana H.Z Robinson[2*], MPH; Laura Harker[2], BS-MPH (Current); Kimberly R. Jacob Arriola[2], MPH, PhD

[1]Emory University School of Medicine, Atlanta Clinical and Translational Science Institute, Emory University, Atlanta, GA, United States

[2]Rollins School of Public Health, Department of Behavioral Sciences and Health Education, Emory University, Atlanta, GA, United States

[*]these authors contributed equally

Corresponding Author:
Dana H.Z Robinson, MPH
Rollins School of Public Health
Department of Behavioral Sciences and Health Education
Emory University
1518 Clifton Rd., NE
Rm 525
Atlanta, GA, 30322
United States
Phone: 1 404 594 1156
Fax: 1 404 727 1369
Email: dhrobin@emory.edu

Abstract

The use of Web-based methods for research recruitment and intervention delivery has greatly increased as Internet usage continues to grow. These Internet-based strategies allow for researchers to quickly reach more people. African-Americans are underrepresented in health research studies. Due to this, African-Americans get less benefit from important research that could address the disproportionate health outcomes they face. Web-based research studies are one promising way to engage more African-Americans and build trust with the African-American community. With African-Americans' increasing access to the Internet using mobile phones and other mobile phone technologies, we advocate for efforts to increase the representation of African-Americans in research studies by using the Internet as a recruitment tool and conclude with recommendations that support this goal.

Introduction

African-Americans are commonly underrepresented in research studies [1-5]. A myriad of concerns are commonly cited when describing the reasons for this underrepresentation and why African-Americans are reluctant to participate in research studies [6-15]. Central to this lack of willingness to take part in research is an overarching sense of distrust that stems from a history of medical injustices [11,16,17]. Although the Tuskegee Syphilis Study is often cited as the most salient historical example of injustice in research pertaining to African-Americans, there are many other examples of medical racism in African-Americans' collective consciousness [18-20]. Additional apprehensions that African-Americans tend to report are often accompanied by a lack of understanding related to the importance of research and

the research process, economic challenges (transportation, employment) related to participation [21,22], and inadequate recruitment efforts made by study investigators [21-23]. Despite the challenges with research participation, increased morbidity and mortality related to chronic disease, predisposition to certain health conditions, and disproportionate impact of illness and varying health outcomes [24,25] necessitate concerted effort for engaging this population in health research. The research literature provides limited guidance in terms of effective strategies for using technology as a recruitment platform for African-Americans. However, such strategies are needed to better harness the benefits of technology and engage African-American participation in health research.

Existing studies have commonly recruited African-Americans via religious institutions and affiliations, community networks

and organizations, and one-on-one interaction [26]. There are several specific strategies that have demonstrated effectiveness in the past. These strategies entail (1) direct face-to-face interaction with individuals and groups; (2) the necessity for flexibility with the manner in which participants are contacted (mobile phone, in-person, email); (3) provision of varying forms of information (booklet, handout, flier); (4) tangible participant incentives; and (5) an in-person reassurance of comfort with the research process inclusive of open sharing of information, opportunity for questions, and assertion of confidentiality [21]. However, when research necessitates large and diverse samples, many of these strategies are time consuming and often cost-prohibitive. Internet-based studies, in particular may require a different approach to recruitment, given that African-Americans' underrepresentation in Web-based studies is well documented [21,27-32]. The purpose of this paper was to argue for enhanced efforts to include African-American participants in Internet-based research studies, despite the many challenges with doing so, and offer practical recommendations for effective recruitment strategies.

Relevant Theoretical and Ethical Considerations

There are important theoretical and ethical considerations that drive decisions about, and describe the importance of, how to engage African-Americans in Internet-based research. Childers and Skinner developed equity theory as a useful framework for understanding the processes that maximize participation in survey research; and we believe that its application within a Web-based research context has far-reaching implications [33]. The authors argue that in a research study, the researcher must establish trust with the participant if the expectation is for the participant to comply with the protocol for the study. Equity theory proposes that when a researcher and participant interact, there is a comparison of the input from the participant and the gains as a result of that input [33]. Thus, an effective interaction (exchange) would be reflected if the participant feels that his or her input is equal to the outcome or gain. The subsequent reward (outcome) can be monetary, in terms of an incentive in exchange for participating, or psychological, reflecting a symmetry with the research topic and his or her values [33]. In light of studies finding greater distrust of medical research among African-Americans than other groups [34-36], there is greater responsibility on behalf of researchers to ensure that participants perceive an equitable exchange with their research participation. For example, the recruitment and screening process could be structured so that a study staff member personally explains the role (input) and benefits (ie, incentive) of participation in lieu of a completely automated study that requires no interaction with study staff for participation. An explanation of this sort provides participants with an opportunity for a personal cost–benefit analysis for research participation that ideally tips the scales in favor of participating.

In addition to equity theory, one can think about African-Americans' involvement in Web-based research through a research ethics lens. Health researchers are trained to abide by 3 ethical principles: respect for persons, beneficence, and justice [37]. From an ethical perspective, all research participants should be treated as autonomous individuals who are able to make their own decisions, and persons with diminished autonomy are protected. Under the principle of beneficence, harm to participants is minimized, and possible benefits of the research are maximized. However, the principle of justice is especially relevant to the current topic because it requires equity among research subjects and equal distribution of benefits across populations. This principle lays foundation for claims that a proportionate number of African-Americans participating in health research is necessary so that results can be generalized to African-Americans, and their health can be impacted by medical advances. With an increasing volume of Internet-based research, the opportunity for participation among African-Americans has expanded giving rise to greater opportunity for study participation and subsequent benefit of health research [38-43].

Internet Research Advantages

Using the Internet for health research has several main advantages. First, recruiting participants through the Internet allows for faster data collection at a lower cost than traditional in-person and email-based methods [44]. Participants can be screened almost instantly, and information can be entered directly into Web-based databases, thus reducing the time-intensive manual data entry process. Second, Web-based recruitment has the potential for farther reach unlike traditional in-person recruitment, which is bound by geographic restrictions [45]. This is especially beneficial when the target population is relatively small, homebound, or lacks transportation and when studying a sensitive topic [14]. Given that 87% of people in the United States are estimated to have access to the Internet, the Internet can be a powerful tool for reaching these groups [46]. Internet use among young, college-educated people with higher incomes is comparable regardless of race and across Internet platforms, African-Americans and whites are very similar with mobile Internet access [47]. Mobile phones play a significant role bridging this Internet access gap [48]. Third, Web-based methods can eliminate common participant barriers to participation, such as lack of time, transportation, and scheduling [25]. Participating in research through the Internet tends to be more convenient as the participant can usually complete the research within his or her own time frame and location of preference [49].

Social media is one Web-based method that can be used to recruit specific individuals based on a specific demographic profile and ensure access to a diverse group of people [29]. For example, a 2014 study used social media sites to recruit HIV-positive individuals. Overall, 1221/1404 (87%) eligible participants completed the survey, indicating that Web-based recruitment was effective [49]. Using self-reported information on social media user profiles, researchers can target advertisements for specific populations and change content. For example, in a study conducted by Sullivan et al. participants were more likely to respond if the people pictured in the advertisements belonged to their racial or ethnic group [29].

Social media sites are also increasingly more popular among African-Americans. Research studies that use social media sites can especially benefit from recruiting a younger African-American audience, as this population uses social media at rates higher than comparable figures for whites [47]. Twitter specifically has a larger proportion of African-American users than whites based on results from a Pew Research study using a nationally representative sample (117/532, 22% of African-Americans; 579/3617, 16% of whites) [47]. Research has demonstrated that targeted ads on social media sites are effective at recruiting specialized populations [44,50]. For example, a 2013 Web-based preventive depression study recruited participants using search engine advertising (Google, Yahoo!, Bing), Facebook advertising, posts in forums and Web-based noticeboards, and promotion through relevant websites and email newsletters of mental health organizations. Several of the methods were effective, but Google Ads yielded the most participants [45]. Increasingly, researchers have used a variety of Web-based methods to recruit a broad sample of participants such as Web-based advertisements, Craigslist, and mobile phone apps [51].

Internet Research Disadvantages

Although the digital divide has been closing over time and African-Americans represent a group adopting broadband Internet at rates greater than other ethnic groups, there are still gaps in Internet access in the United States [52]. Large disparities exist with regard to the location, frequency, duration, and type of Internet access; however, this gap is not consistent across technology platforms [47]. As of 2014, using a nationally representative sample, it was estimated that 531/664 (80%) of African-Americans had Internet access (compared with 3674/4223, 87% of whites), and 412/664 (62%) had a home broadband connection (compared with 3125/4223, 74% of whites) [47]. Internet usage rates are most notable when comparing differences between African-Americans (299/664, 45%) and whites (2660/4223, 63%) with respect to older adults and those with lower levels of education [47]. Consequently, persons with less access to technology use email and social networking apps less often than those with more consistent access. These differences can affect completion rates for Internet surveys and are likely confounded by education and literacy [26]. The differences in completion rates can then create bias within the sample, which affects the generalizability of study findings.

General challenges with Web-based research participation are also abundant and entail issues related to recruitment, response rate, study retention, generalizability, and overall study design [14,15]. Some common reasons for low response rates among potential participants include ignoring study emails or ads, lacking the motivation (e.g., incentives, personal connection to the study) to participate, and/or difficulty with distinguishing legitimate messages from spam [53]. Minority recruitment can be even more complicated and difficult when compared with recruitment of other populations because of the inherent characteristics of the Internet such as concerns with absolute authenticity, the necessity of flexibility with multiple recruitment strategies, and dependency on cooperation with gatekeepers [15]. These are all challenges that should be accounted for when considering the Internet as a tool for research studies targeting African-Americans.

Recommendations

There is value in conducting Internet-based research studies targeting African-Americans to improve their health and well-being. Oftentimes, research with African-Americans requires a level of sensitivity that recognizes the influence of culture, history, and familial interactions. As a result, more efforts must be made by researchers who intend to include African-Americans in ways that align with equity theory [33].

We offer 5 recommendations to increase the successful recruitment of African-Americans into Internet-based research. First, consider using recruiters. In this approach, African-Americans are trained (via webinar) to recruit from their social networks both through the Internet and in person. The training should involve a brief overview of the reasons for African-Americans' hesitancy to be involved in research studies (as we described previously), the importance of justice and equity theory as it relates to African-American recruitment, background information on the study, and techniques for Web-based recruitment such as social media. By using social media, trust is already established between the researcher (via the recruiter) and potential participants, increasing the likelihood of enrolling participants into the study. In addition to the potential for a large reach, social media recruitment has been demonstrated to be cost-effective when compared with traditional recruitment methods [43,54,55]. Second, and related to the first, is to capitalize from a commonly used technique—snowball sampling—in which participants recruit other participants [56]. Respondent-driven sampling is a commonly used form of snowball sampling [57]. These methods are being refined for use in a Web-based environment but may be particularly helpful for recruiting African-Americans because of previously established trusted relationships that exist among participants and members of their social networks.

Third, consider offering a list of incentives from which participants can choose. Each incentive must have the ability to be delivered quickly and with minimal effort through the Internet (eg, electronic methods for cash payment such as PayPal, electronic gift cards such as Amazon e-gift cards, and virtual debit or credit cards). An incentive that is worthwhile to the participant will cause him or her to feel that the exchange is equitable. That is, the participant perceives that he or she is receiving a reward equal to his or her effort. Fourth, 2-tiered recruitment is suggested when feasible (eg, there is a small geographic recruitment area). In this approach, participant recruitment occurs initially in person affording the participant an opportunity to interact with a trained peer volunteer and establish a rapport. Simultaneously, the participant will receive Web-based communication with directions for the completion of an Internet-based screener. The second tier of this approach entails the participant completing the screener, intervention components, and all questionnaires online. Decreased reach is a limitation of this approach because a researcher working face to face with participants cannot engage as many participants as

he or she could do through the Internet. A final recommendation is to use trusted institutions as partners for participant recruitment. Institutions such as churches, community organizations, sororities, and fraternities have existing relationships with the communities of which they are a part. The investigator can establish relationships with members of the institutions directly then work with the institution to send emails to their membership regarding study participation.

Conclusions

The convenient, far-reaching, fast-paced, Web-based environment provides numerous methodological advantages for researchers, yet specialized efforts are necessary to ensure the inclusion of African-American participants in Internet-based research studies. To build trust within African-American communities, address the unique challenges posed by participation in research. The hesitancy related to participation, is often a combination of a lack of understanding related to the relevancy of research [21,22] inadequate recruitment efforts made by study investigators [21-23], and a general lack of trust with respect to the health and medical field [11,16]. This viewpoint recommended 5 approaches that entail using recruiters and previous participants themselves, multiple-choice incentives, a 2-tiered recruitment strategy, and partnership with trusted organizations. These approaches emphasize the importance of relationship building as this has been highlighted as a key component of research recruitment and retention of African-American participants [21,58]. The use of these approaches may not be particular to African-Americans; however, they may be particularly useful for investigators to overcome some of the challenges related to mistrust and serve as a mechanism to establish a greater rapport with African-American research participants.

Conflicts of Interest

None declared.

References

1. Zahm S, Pottern L, Lewis D, Ward M, White D. Inclusion of women and minorities in occupational cancer epidemiologic research. J Occup Med 1994 Aug;36(8):842-847. [Medline: 7807263]
2. Svensson C. Representation of American blacks in clinical trials of new drugs. JAMA 1989 Jan 13;261(2):263-265. [Medline: 2909024]
3. Heiat A, Gross C, Krumholz H. Representation of the elderly, women, and minorities in heart failure clinical trials. Arch Intern Med 2002;162(15):1682-1688. [Medline: 12153370]
4. Hall W. Representation of blacks, women, and the very elderly (aged > or = 80) in 28 major randomized clinical trials. Ethn Dis 1999;9(3):333-340. [Medline: 10600055]
5. Ashing-Giwa K, Ganz P, Petersen L. Quality of life of African-American and white long term breast carcinoma survivors. Cancer 1999 Jan 15;85(2):418-426. [Medline: 10023710]
6. Jacob AK, Perryman J, Doldren M. Moving beyond attitudinal barriers: understanding African Americans' support for organ and tissue donation. J Natl Med Assoc 2005 Mar;97(3):339-350. [Medline: 15779498]
7. Durand R, Decker PJ, Bruder P. Organ donation among African Americans: opportunities for increasing donor rates. Hosp Top 2002;80(3):34-37. [doi: 10.1080/00185860209598002] [Medline: 12471884]
8. Minniefield WJ, Yang J, Muti P. Differences in attitudes toward organ donation among African Americans and whites in the United States. J Natl Med Assoc 2001 Oct;93(10):372-379. [Medline: 11688917]
9. Morgan S, Miller J, Arasaratnam L. Similarities and Differences Between African Americans' and European Americans' Attitudes, Knowledge, and Willingness to Communicate About Organ Donation1. J Appl Social Pyschol 2003 Apr;33(4):693-715. [doi: 10.1111/j.1559-1816.2003.tb01920.x]
10. Terrell F, Moseley KL, Mosley KL, Terrell AS, Nickerson KJ. The relationship between motivation to volunteer, gender, cultural mistrust, and willingness to donate organs among Blacks. J Natl Med Assoc 2004 Jan;96(1):53-60. [Medline: 14746354]
11. Blanchard J, Lurie N. R-E-S-P-E-C-T: patient reports of disrespect in the health care setting and its impact on care. J Fam Pract 2004 Sep;53(9):721-730. [Medline: 15353162]
12. Carpenter WR, Godley PA, Clark JA, Talcott JA, Finnegan T, Mishel M, et al. Racial differences in trust and regular source of patient care and the implications for prostate cancer screening use. Cancer 2009 Nov 1;115(21):5048-5059 [FREE Full text] [doi: 10.1002/cncr.24539] [Medline: 19637357]
13. Mechanic D, Meyer S. Concepts of trust among patients with serious illness. Soc Sci Med 2000 Sep;51(5):657-668. [Medline: 10975226]
14. Hamilton RJ, Bowers BJ. Internet recruitment and e-mail interviews in qualitative studies. Qual Health Res 2006 Jul;16(6):821-835. [doi: 10.1177/1049732306287599] [Medline: 16760538]

15. Im E, Chee W. Methodological issues in the recruitment of ethnic minority subjects to research via the Internet: a discussion paper. Int J Nurs Stud 2005 Nov;42(8):923-929. [doi: 10.1016/j.ijnurstu.2005.01.002] [Medline: 16210030]

16. Brandon DT, Isaac LA, LaVeist TA. The legacy of Tuskegee and trust in medical care: is Tuskegee responsible for race differences in mistrust of medical care? J Natl Med Assoc 2005 Jul;97(7):951-956. [Medline: 16080664]

17. Gamble V. Under the shadow of Tuskegee: African Americans and health care. Am J Public Health 1997 Nov;87(11):1773-1778. [Medline: 9366634]

18. Skloot R. The Immortal Life of Henrietta Lacks. New York: Crown Publishers; 2010.

19. Washington H. Medical apartheid: the dark history of medical experimentation on Black Americans from colonial times to the present. New York: Doubleday; 2006.

20. Randall VR. Slavery, segregation and racism: trusting the health care system ain't always easy! An African American perspective on bioethics. St Louis Univ Public Law Rev 1996;15(2):191-235. [Medline: 11656870]

21. Qualls CD. Recruitment of African American adults as research participants for a language in aging study: example of a principled, creative, and culture-based approach. J Allied Health 2002;31(4):241-246. [Medline: 12491954]

22. Coker A, Huang H, Kashubeck-West S. Research with African Americans: Lessons learned about recruiting African American women. Journal of Multicultural Counseling and Development 2009;37(3):153-165. [doi: 10.1002/j.2161-1912.2009.tb00099.x]

23. Williams I, Utz SW, Jones R, Hinton I, Steeves R, Alexander G. Recruitment of Rural African Americans for Research Projects: Lessons Learned. South Online J Nurs Res 2011 Apr 1;11(1):8 [FREE Full text] [Medline: 24791157]

24. Centers for Disease Control and Prevention. Atlanta: Centers for Disease Control and Prevention, National Center for Health Statistics; 2016. Health of Black or African American non-Hispanic Population URL: http://www.cdc.gov/nchs/fastats/black-health.htm [accessed 2016-06-02] [WebCite Cache ID 6hy2XtAgG]

25. Branson RD, Davis K, Butler KL. African Americans' participation in clinical research: importance, barriers, and solutions. Am J Surg 2007 Jan;193(1):32-9; discussion 40. [doi: 10.1016/j.amjsurg.2005.11.007] [Medline: 17188084]

26. Satia JA, Galanko JA, Rimer BK. Methods and strategies to recruit African Americans into cancer prevention surveillance studies. Cancer Epidemiol Biomarkers Prev 2005 Mar;14(3):718-721 [FREE Full text] [doi: 10.1158/1055-9965.EPI-04-0132] [Medline: 15767356]

27. Khosropour CM, Sullivan PS. Predictors of retention in an online follow-up study of men who have sex with men. J Med Internet Res 2011;13(3):e47 [FREE Full text] [doi: 10.2196/jmir.1717] [Medline: 21745792]

28. Bender M, Clark M, Guevara E, Chee W, Im E. Challenges in internet study recruitment of African American cancer patients. Stud Health Technol Inform 2006;122:878-879. [Medline: 17102438]

29. Sullivan PS, Khosropour CM, Luisi N, Amsden M, Coggia T, Wingood GM, et al. Bias in online recruitment and retention of racial and ethnic minority men who have sex with men. J Med Internet Res 2011;13(2):e38 [FREE Full text] [doi: 10.2196/jmir.1797] [Medline: 21571632]

30. Whittemore R, Jaser S, Faulkner M, Murphy K, Delamater A, Grey M. Type 1 diabetes eHealth psychoeducation: youth recruitment, participation, and satisfaction. J Med Internet Res 2013;15(1):e15 [FREE Full text] [doi: 10.2196/jmir.2170] [Medline: 23360729]

31. Green BB, Anderson ML, Ralston JD, Catz S, Fishman PA, Cook AJ. Patient ability and willingness to participate in a web-based intervention to improve hypertension control. J Med Internet Res 2011;13(1):e1 [FREE Full text] [doi: 10.2196/jmir.1625] [Medline: 21371993]

32. Smith YR, Johnson AM, Newman LA, Greene A, Johnson Timothy R B, Rogers JL. Perceptions of clinical research participation among African American women. J Womens Health (Larchmt) 2007 Apr;16(3):423-428 [FREE Full text] [doi: 10.1089/jwh.2006.0124] [Medline: 17439387]

33. Childers T, Skinner S. Toward a conceptualization of mail survey response behavior. Psychol. Mark 1996 Feb;13(2):185-209. [doi: 10.1002/(SICI)1520-6793(199602)13:2<185::AID-MAR5>3.0.CO;2-B]

34. Corbie-Smith G, Thomas SB, Williams MV, Moody-Ayers S. Attitudes and beliefs of African Americans toward participation in medical research. J Gen Intern Med 1999 Sep;14(9):537-546 [FREE Full text] [Medline: 10491242]

35. LaVeist TA, Nickerson KJ, Bowie JV. Attitudes about racism, medical mistrust, and satisfaction with care among African American and white cardiac patients. Med Care Res Rev 2000;57 Suppl 1:146-161. [Medline: 11092161]

36. Shavers VL, Lynch CF, Burmeister LF. Factors that influence African-Americans' willingness to participate in medical research studies. Cancer 2001 Jan 1;91(1 Suppl):233-236. [Medline: 11148585]

37. US Department of Health and Human Services. Office of Human Research Protections. Washington, DC; 1979. The Belmont Report URL: http://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html [accessed 2016-06-02] [WebCite Cache ID 6hy4VsMI1]

38. Clarke G, Reid E, Eubanks D, O'Connor E, DeBar LL, Kelleher C, et al. Overcoming depression on the Internet (ODIN): a randomized controlled trial of an Internet depression skills intervention program. J Med Internet Res 2002 Dec;4(3):E14 [FREE Full text] [doi: 10.2196/jmir.4.3.e14] [Medline: 12554545]

39. Tate DF, Wing RR, Winett RA. Using Internet technology to deliver a behavioral weight loss program. JAMA 2001 Mar 7;285(9):1172-1177. [Medline: 11231746]

40.   Vinokur AD, Merion RM, Couper MP, Jones EG, Dong Y. Educational web-based intervention for high school students to increase knowledge and promote positive attitudes toward organ donation. Health Educ Behav 2006 Dec;33(6):773-786. [doi: 10.1177/1090198106288596] [Medline: 16923836]

41.   Ruggiero L, Moadsiri A, Quinn L, Riley BB, Danielson KK, Monahan C, et al. Diabetes island: preliminary impact of a virtual world self-care educational intervention for african americans with type 2 diabetes. JMIR Serious Games 2014;2(2):1-12 [FREE Full text] [doi: 10.2196/games.3260] [Medline: 25584346]

42.   Merion RM, Vinokur AD, Couper MP, Jones EG, Dong Y, Wimsatt M, et al. Internet-based intervention to promote organ donor registry participation and family notification. Transplantation 2003 Apr 27;75(8):1175-1179. [doi: 10.1097/01.TP.0000062845.51042.DD] [Medline: 12717199]

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

44.   King D, O'Rourke N, DeLongis A. Social media recruitment and online data collection: A beginner's guide and best practices for accessing low-prevalence and hard-to-reach populations. Canadian Psychology/Psychologie canadienne 2014;55(4):240-249. [doi: 10.1037/a0038087]

45.   Morgan A, Jorm A, Mackinnon A. Internet-based recruitment to a depression prevention intervention: lessons from the Mood Memos study. J Med Internet Res 2013;15(2):e31 [FREE Full text] [doi: 10.2196/jmir.2262] [Medline: 23403043]

46.   The World Bank. Washington, DC: World Bank Group; 2015. Internet Users (per 100 people) URL: http://data.worldbank.org/indicator/IT.NET.USER.P2 [accessed 2016-06-02] [WebCite Cache ID 6hy5SIFdN]

47.   Smith A. PEW Research Center: Internet Science and Tech. Washington, D.C: PEW Research Center; 2014. African Americans and Technology Use URL: http://www.pewinternet.org/2014/01/06/african-americans-and-technology-use/ [accessed 2016-06-02] [WebCite Cache ID 6hy5djpRG]

48.   The Nielsen Company.: Nielsen; 2015. Multifaceted connections: African Americans media usage outpaces across platforms URL: http://www.nielsen.com/us/en/insights/news/2015/multifaceted-connections-african-american-media-usage-outpaces-across-platforms.html [accessed 2016-06-02] [WebCite Cache ID 6hy5r31cX]

49.   Yuan P, Bare M, Johnson M, Saberi P. Using online social media for recruitment of human immunodeficiency virus-positive participants: a cross-sectional survey. J Med Internet Res 2014;16(5):e117 [FREE Full text] [doi: 10.2196/jmir.3229] [Medline: 24784982]

50.   Fenner Y, Garland S, Moore E, Jayasinghe Y, Fletcher A, Tabrizi SN, et al. Web-based recruiting for health research using a social networking site: an exploratory study. J Med Internet Res 2012;14(1):e20 [FREE Full text] [doi: 10.2196/jmir.1978] [Medline: 22297093]

51.   Worthen M. An invitation to use craigslist ads to recruit respondents from stigmatized groups for qualitative interviews. Qualitative Research 2013 Apr 02;14(3):371-383. [doi: 10.1177/1468794113481791]

52.   Zickuhr K, Smith A. Pew Research Center: Internet, Science & Tech. Washington, DC: Pew Research Center; 2013. Home Broadband 2013 URL: http://www.pewinternet.org/2013/08/26/home-broadband-2013/ [accessed 2016-06-02] [WebCite Cache ID 6hy6b9cEH]

53.   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]

54.   Valdez R, Guterbock T, Thompson M, Reilly JD, Menefee HK, Bennici MS, et al. Beyond traditional advertisements: leveraging Facebook's social structures for research recruitment. J Med Internet Res 2014;16(10):e243 [FREE Full text] [doi: 10.2196/jmir.3786] [Medline: 25348050]

55.   Gross MS, Liu NH, Contreras O, Muñoz RF, Leykin Y. Using Google AdWords for international multilingual recruitment to health research websites. J Med Internet Res 2014;16(1):e18 [FREE Full text] [doi: 10.2196/jmir.2986] [Medline: 24446166]

56.   Baltar F, Brunet I. Social research 2.0: virtual snowball sampling method using Facebook. Internet Research 2012 Jan 27;22(1):57-74. [doi: 10.1108/10662241211199960]

57.   Heckathorn D. Respondent-driven sampling: a new approach to the study of hidden populations. Social Problems 1997 May;44(2):174-199.

58.   Mason SE. Offering African Americans opportunities to participate in clinical trials research: how social workers can help. Health Soc Work 2005 Nov;30(4):296-304. [Medline: 16323721]

XSL•FO
RenderX

<u>Original Paper</u>

Web-Based Fully Automated Self-Help With Different Levels of Therapist Support for Individuals With Eating Disorder Symptoms: A Randomized Controlled Trial

Jiska J Aardoom[1], MSc; Alexandra E Dingemans[1], PhD; Philip Spinhoven[2,3], PhD; Joost R van Ginkel[4], PhD; Mark de Rooij[2], PhD; Eric F van Furth[1,3], PhD

[1]Rivierduinen Eating Disorders Ursula, Leiden, Netherlands

[2]Institute of Psychology, Leiden University, Leiden, Netherlands

[3]Leiden University Medical Center, Department of Psychiatry, Leiden, Netherlands

[4]Institute of Education and Child Studies, Leiden University, Leiden, Netherlands

Corresponding Author:
Jiska J Aardoom, MSc
Rivierduinen Eating Disorders Ursula
PO BOX 405
Leiden, 2300 AK
Netherlands
Phone: 31 718903009
Fax: 31 718903009
Email: j.aardoom@rivierduinen.nl

Abstract

Background: Despite the disabling nature of eating disorders (EDs), many individuals with ED symptoms do not receive appropriate mental health care. Internet-based interventions have potential to reduce the unmet needs by providing easily accessible health care services.

Objective: This study aimed to investigate the effectiveness of an Internet-based intervention for individuals with ED symptoms, called "Featback." In addition, the added value of different intensities of therapist support was investigated.

Methods: Participants (N=354) were aged 16 years or older with self-reported ED symptoms, including symptoms of anorexia nervosa, bulimia nervosa, and binge eating disorder. Participants were recruited via the website of Featback and the website of a Dutch pro-recovery–focused e-community for young women with ED problems. Participants were randomized to: (1) Featback, consisting of psychoeducation and a fully automated self-monitoring and feedback system, (2) Featback supplemented with low-intensity (weekly) digital therapist support, (3) Featback supplemented with high-intensity (3 times a week) digital therapist support, and (4) a waiting list control condition. Internet-administered self-report questionnaires were completed at baseline, post-intervention (ie, 8 weeks after baseline), and at 3- and 6-month follow-up. The primary outcome measure was ED psychopathology. Secondary outcome measures were symptoms of depression and anxiety, perseverative thinking, and ED-related quality of life. Statistical analyses were conducted according to an intent-to-treat approach using linear mixed models.

Results: The 3 Featback conditions were superior to a waiting list in reducing bulimic psychopathology (d=−0.16, 95% confidence interval (CI)=−0.31 to −0.01), symptoms of depression and anxiety (d=−0.28, 95% CI=−0.45 to −0.11), and perseverative thinking (d=−0.28, 95% CI=−0.45 to −0.11). No added value of therapist support was found in terms of symptom reduction although participants who received therapist support were significantly more satisfied with the intervention than those who did not receive supplemental therapist support. No significant differences between the Featback conditions supplemented with low- and high-intensity therapist support were found regarding the effectiveness and satisfaction with the intervention.

Conclusions: The fully automated Internet-based self-monitoring and feedback intervention Featback was effective in reducing ED and comorbid psychopathology. Supplemental therapist support enhanced satisfaction with the intervention but did not increase its effectiveness. Automated interventions such as Featback can provide widely disseminable and easily accessible care. Such interventions could be incorporated within a stepped-care approach in the treatment of EDs and help to bridge the gap between mental disorders and mental health care services.

XSL•FO

RenderX

Introduction

Eating disorders (EDs) are serious psychiatric disorders characterized by high rates of comorbidity, chronicity, mortality, and relapse [1-5]. Unfortunately, despite the disabling nature of these disorders, many individuals with ED symptoms do not seek and receive appropriate mental health care [2,6]. Barriers to care include geographical or financial barriers, as well as fear of stigmatization and feelings of shame [7]. E-mental health has the potential to reduce these barriers in help-seeking, as well as the unmet need for health care by providing easily accessible services.

Numerous Internet-based interventions for the prevention and treatment of ED have shown promising results [8-10]. The results of a recent meta-analytic review [10] demonstrated that Internet-based programs, of which the majority was based on cognitive behavioral principles, were successful in decreasing a range of ED-related symptoms including body dissatisfaction, symptoms of bulimia nervosa, shape and weight concerns, dietary restriction, and negative affect. Emerging research furthermore suggests that ehealth interventions may reach underserved populations and increase access to regular health care [11]. Despite the promising results, research into the effectiveness of such interventions is still in an early stage [9,12,13], and further high-quality studies are required.

Internet-based interventions can include many different components and can be provided with or without therapist support. In the field of depression and anxiety, it has been found that Internet-based interventions with therapist support were more effective than those without or those with only minimal therapeutic contact [14,15]. Direct comparisons of Internet-based mental health interventions with and without therapist support in randomized controlled trials are scarce although a recent meta-analysis indeed demonstrated guided interventions to be superior to unguided interventions [16]. However, studies investigating the optimal intensity of therapist support are rare [16], and it is currently unknown how much or how little therapist support is needed to realize a particular amount of additional improvement in health outcomes. To our knowledge, only 1 study directly compared different intensities of therapist support in an Internet-based treatment for panic disorder [17]. This study demonstrated no significant differences between higher and lower intensities of therapist support. Regarding ehealth interventions in the field of ED, no studies have yet directly compared guided and nonguided interventions nor have different intensities of therapist support been investigated.

In addition to the intensity of therapist support, another important factor is the way in which such support is provided. Tate et al [18] investigated the effectiveness of feedback on

self-monitoring diaries provided by either a human counselor or a computer-automated program in an Internet-based weight loss program. Interestingly, at 3-month follow-up, no significant differences in outcome were found between participants in the computer-automated counseling condition and the human counseling condition, respectively. Along similar lines, a recent study demonstrated a Web-based intervention for mild-to-moderate depression symptoms to be equally effective when provided with human versus automated support [19]. Hence, automated support may be an effective and widely disseminable means of providing support within Internet-based interventions, and it is important to further compare the effectiveness of such automated support with the effectiveness of different intensities of individual therapist support.

This study evaluated self-help intervention "Featback" for individuals with ED symptoms. Featback comprises psychoeducation and a fully automated self-monitoring and feedback system. Self-monitoring is an important clinical technique that is often used in cognitive behavioral therapy [20], where it can among other things help to gain a more comprehensive understanding of one's psychopathology. By means of the monitoring and feedback system, participants are invited to complete a weekly monitoring questionnaire assessing the core symptoms of ED: body dissatisfaction, excessive concern with body weight and shape, unbalanced nutrition and dieting, and binge eating and compensatory behaviors. After completion of the questionnaire, participants receive a feedback message, which is automatically generated and tailored to their answers of the monitoring questions, containing social support and advice on how to counteract reported ED symptoms. Featback is aimed at individuals with all types of ED symptoms, which in line with the transdiagnostic theory that all EDs (eg, anorexia nervosa, bulimia nervosa, binge eating disorder) share the same core psychopathology, characterized by the overevaluation of eating, shape, weight, and their control [21].

The first aim of this study was to investigate the effectiveness of Featback in reducing ED psychopathology and comorbid symptoms. The second aim was to investigate the added value of therapist support and different intensities of therapist support. A randomized controlled trial was conducted comparing 4 conditions: (1) Internet-based intervention Featback, consisting of psychoeducation and a fully automated monitoring and feedback system, (2) Featback supplemented with low-intensity (weekly) therapist support, (3) Featback supplemented with high-intensity (3 times a week) therapist support, and (4) a waiting list control (WLC).

Methods

Study Design and Procedure

This study was a 4-arm randomized control trial. Ethical approval was obtained from the Leiden University Medical Center Ethics Committee. This committee granted exemption for parental consent for individuals aged between 16 and 18 years of age. Detailed information on the study methods, including the design, intervention conditions, measures, and ethical precautions and crisis management, can be found in the published study protocol [22].

Participants were recruited via the website of Featback [23] and the website of Dutch pro-recovery–focused e-community "Proud2Bme" [24] for young women with ED problems. The eligibility criteria were: (1) age ≥ 16 years, (2) access to the Internet, and (3) ED symptoms. The latter was defined as scoring ≥ 52 on the Weight Concern Scale [25] or reporting 1 or more of the following ED symptoms as assessed by the Short Evaluation of Eating Disorders (SEED) [26]: a body mass index of ≤18.5, ≥ 1 binge eating episodes a week over the past 4 weeks, and engagement in ≥ 1 compensatory behaviors a week over the past 4 weeks.

After Web-based completion of informed consent and the screening questionnaire including questions regarding the eligibility criteria, participants were invited to complete the baseline questionnaire. Thereafter, participants were randomly assigned to 1 of the 4 study conditions with a block size of 40 and an equal allocation ratio (1:1:1:1). An independent researcher who had no involvement in any other aspect of this study conducted the randomization allocation by means of computer-generated random numbers created in SPSS. She concealed the allocation sequence in a password-protected computer file from the main researchers until interventions were assigned, preventing researchers from having any prior knowledge of the upcoming condition assignments. Importantly, therapists were alternately assigned to low- versus high-intensity therapist support.

Interventions

Featback

All participants had access to the Featback website where comprehensive and general information on ED could be found (ie, psychoeducation), for example, the types of EDs and symptoms, risk factors, causes, and comorbid problems. This information served primarily to educate participants about EDs and stimulate recognition and acknowledgement. The psychoeducation was purely self-guided, meaning that participants were free in choosing when and what to read. The monitoring and feedback system comprised a weekly invitation by email to complete a monitoring questionnaire. This questionnaire consisted of 8 4-point Likert items assessing cognitive and behavioral correlates of the following 4 dimensions: (1) body dissatisfaction, (2) excessive concerns with body weight and shape, (3) unbalanced nutrition and dieting, and (4) binge eating and compensatory behaviors. After completion, an algorithm determines the patterns of change of each of these 4 dimensions: still in the functional or healthy

range, still in the dysfunctional or unhealthy range, improvement from the dysfunctional to the functional range, or deterioration from the functional to the dysfunctional range. The 4 different patterns of change with respect to the 4 dimensions of ED symptoms result in 4×4×4×4= 256 possible scenarios regarding a participant's status. For each possible scenario, 10 to 15 different feedback messages were preformulated in a database. After determining the status of a participant, the algorithm randomly selected 1 tailored feedback message out of this database and sent this to the participant accordingly. Hence, when a participant's status does not change over time, one would not receive the same message over and over again. All the feedback messages contained social support by expressing interest in and concerns about the participants' well-being. Positive reinforcement techniques such as encouragement were used to stimulate and maintain healthy behaviors and attitudes. Furthermore, the messages included tips and advice on how to counteract negative developments in reported ED-related symptoms. The following is an example of a feedback message, which could be sent to someone with dysfunctional overconcerns with body weight and shape, unbalanced nutrition and dieting (dysfunctional), as well as deteriorations in body dissatisfaction and symptoms of binge eating and compensatory behaviors:

> *We are concerned with the changes in your body image and eating behaviors, however, we know that you have the ability to make healthy changes. Your body image and eating habits are closely linked. This week, try to eat regular, well-balanced meals and snacks, which might help to prevent the binge eating and/or compensatory behaviors and help you to feel better. If you continue to have negative thoughts about your body, it may be helpful for you to talk to someone about it, maybe a family member? Or a friend? Take care!*

The fully automated self-monitoring and feedback system was developed in Germany, and for more detailed information on this system, see the study by Bauer et al [27]. A reminder was sent to participants by email each time they failed to complete a monitoring assessment.

Featback + Low-Intensity Therapist Support

Participants received Featback as described previously supplemented with low-intensity (weekly) therapist support by means of email, chat and/or audio teleconference (ie, Skype). Participants could schedule support sessions in a Web-based agenda where available time slots of the therapist were presented. For each support session, participants could choose their preferred medium of support. Therapists were instructed to send an email to participants in case they did not schedule any support session(s) or in case they did not show up at scheduled support session(s) and to repeat this process twice per nonresponse. Chat and teleconference sessions had a maximum duration of 20 minutes, whereas an email session contained 1 email reply from the therapist to the participant. The therapist support was independent of the monitoring and feedback system. The chat methodology was based on a 5-phase model: (1) a warm welcome, (2) clarifying the question, (3) determining the goal of the conversation, (4) concrete

XSL·FO
RenderX

elaboration of the goal of the conversation, and (5) closing the circle [28]. The email methodology contained 3 phases: (1) extracting the question, (2) formulating an answer, and (3) checking and rereading the message and sending it [28].

Featback + High-Intensity Therapist Support

Participants received Featback, supplemented with high-intensity (3 times a week) therapist support by means of email, chat, and/or teleconference as described previously.

Waiting List Control Condition

Participants were placed on a waiting list for 5 months, after which they were offered Featback with low-intensity therapist support.

In all 4 intervention conditions, participants were free to undergo any other type of intervention or treatment (ie, usual care).

Therapists

The therapists were 7 females who were either Master of Science students in clinical psychology or individuals with a master's degree in clinical psychology. All therapists underwent training in the delivery and methodology of Internet-based support. Furthermore, they received extensive information on EDs and practiced with case material and expert patients (ie, someone who has experienced an ED themselves and has been successful in managing the disorder) before the start of the trial. Monthly face-to-face supervision sessions were organized by the main researcher (JA), a psychologist (MN), and an experienced psychotherapist (EvF) as a matter of routine professional and ethical care, as well as to reinforce adherence to the protocol. In addition, 2 individual supervision sessions were provided to all therapists during their first month. Thereafter, therapists' adherence to the protocol was regularly checked at random, by checking whether the chats and emails included the 5- and 3-phase model, respectively.

Outcomes

All data were collected by means of Internet-administered self-report questionnaires at baseline, post-intervention (8 weeks after baseline), and at 3- and 6-month follow-up. Waiting list participants were offered Featback with low-intensity therapist support after the 3-month follow-up and were not assessed at 6-month follow-up.

The primary outcome measure was ED psychopathology as measured by the SEED [26] and the Eating Disorder Examination Questionnaire (EDE-Q) [29]. The SEED [26] distinguishes between the main symptoms of anorexia nervosa (underweight, fear of weight gain, distortion of body perception) and bulimia nervosa (binge eating, compensatory behaviors, overconcern with body shape and weight). Total severity indexes were calculated for both dimensions. The SEED has demonstrated validity and was shown to be sensitive to symptom change [26]. Regarding the EDE-Q, a global score of ED psychopathology was calculated by summing and averaging 22 7-point Likert items. The EDE-Q has demonstrated reliability and validity [30], and the internal consistency reliability in the current sample was high (Cronbach α=.88). Higher scores on both the SEED (range 0-3) and the EDE-Q (range 0-6) reflect higher ED psychopathology.

Secondary outcome measures included ED-related quality of life as assessed by the ED-Related Quality of Life Questionnaire (ED-QOL), a validated 25-item questionnaire assessing the influence of eating behaviors and body weight in the psychological, physical and cognitive, financial and work- or school-related domain [31]. The ED-QOL demonstrated excellent internal consistency reliability in this study sample (Cronbach α=.92). Higher scores (range 1-5) reflect lower quality of life. Symptoms of depression and anxiety were measured using the 4-item Patient Health Questionnaire (PHQ-4). The PHQ-4 has demonstrated factorial and construct validity [32] and demonstrated good internal consistency reliability in the current sample (Cronbach α=.83). Higher scores (range 0-12) reflect higher symptom severity. Finally, levels of perseverative thinking (ie, worry and rumination) were assessed using the Perseverative Thinking Questionnaire (PTQ) [33]. The PTQ demonstrated good internal consistency and satisfactory stability [33]. The internal consistency reliability in the current sample was excellent (Cronbach α=.95). Higher scores are indicative of higher levels of perseverative thinking (scale 0-4).

Given that participants were free to undergo any other type of intervention, psychological health care service utilization (ie, appointments with a dietitian, social worker, psychologist, psychiatrist, or psychotherapist) was assessed with the Trimbos/iMTA Questionnaire for Costs Associated with Psychiatric Illness: TiC-P) [34]. User satisfaction was assessed with 2 open-ended questions asking participants for their positive and negative feedback, respectively. In addition, participants were asked to rate their satisfaction with the intervention and their satisfaction with their therapist on a 10-point Likert scale ranging from very dissatisfied (score of 1) to very satisfied (score of 10). Finally, 2 open-ended questions assessed the reasons for dropout attrition (ie, not completing study questionnaires) and nonusage attrition (ie, deregistration from the monitoring and feedback system).

Statistical Analyses

All data were analyzed in SPSS version 22 using 2-tailed tests and α=.05. A target sample size of 344 participants was calculated by the software program Power Analysis and Sample Size version 8.0 (2008) to yield 80% power to detect an expected between-group (pooled Featback conditions vs WLC) difference at post-intervention with an effect size of 0.3, α=.05, and an expected dropout rate of 30% (for more details on power calculation, see the paper by Aardoom et al [22]).

Possible differences in baseline characteristics, dropout rates, and participants' experiences were investigated using chi-square tests and analysis of variances. All data were imputed using multiple imputation methods. Multiple imputations using predictive mean matching were conducted in statistical software program R version 3.02. Interactions were taken into account in the imputation procedure [35]. Multiple imputation methods have several advantages over complete-case analyses or single imputation techniques and are therefore highly recommended [36]. For each variable with missing data, the number of predictor variables was determined by the rule of thumb of 15 cases per potential predictor [37]. For example, in case the data

of 300 participants would be available on a specific variable, 300/15=20 predictor variables could be used to predict missing data on this variable. Then, correlations between the outcome variable and all other variables were investigated, so that the variables that correlated the highest with the outcome variable were chosen as predictors for the missing data on the outcome variables. A total of 100 imputed datasets were generated. Results from all imputed datasets were pooled according to Rubin's rules to account for the uncertainty associated with the imputations [38].

The main analyses were conducted using linear mixed models including random intercepts. All analyses were conducted according to the intent-to-treat approach including all participants who underwent randomization. Three statistical models were specified including time and condition contrasts (for details on models and contrast coding, see Multimedia Appendix 1). Model 1 investigated whether the 3 Featback conditions (pooled) led to better outcomes than the WLC. Model 2 compared Featback without therapist support versus the 2 Featback conditions with therapist support (pooled). Model 3 compared Featback with low- versus high-intensity therapist support. Main analyses were repeated controlling for significant baseline differences between the conditions (ie, age, marital status, and duration of ED psychopathology), and number of received psychological health care appointments. The latter was entered as covariate to examine intervention effects over and above usual care. Also, main analyses were repeated for completers of the intervention only, defined as participants who completed at least 5 monitoring questionnaires (Featback without therapist support), plus at least 5 to 13 therapist support sessions (Featback with low- vs high-intensity therapist support, respectively).

Effect sizes (d) were calculated by dividing the unstandardized coefficients of interaction effects (time × condition) by the pooled within-group standard deviation (SD) of the outcome measure at baseline [39]. The resulting effect sizes of all imputed dataset were summed and averaged. The 2 open-ended questions related to satisfaction with the intervention, both critical and positive, were qualitatively explored to provide an overview of participants' most frequently reported negative and positive comments.

Results

Participants

Participants were recruited between November 7, 2012 and June 17, 2013. Follow-up was completed at March 3, 2014. Figure 1 presents the flow of participants through each stage of the trial. A total of 354 participants were assessed at baseline, 273 (77.1%) at post-intervention, 202 (57.1%) at 3-month follow-up, and 118 participants (44.7%) of the available 3 study conditions (n=264) at 6-month follow-up. Study dropout rates did not significantly differ between the conditions at post-intervention ($\chi^2(3)$=4.35, P=.23) and 6-month follow-up ($\chi^2(2)$=2.87, P=.24), although at 3-month follow-up, the WLC participants dropped out of the study less often than participants who received Featback without or with low-intensity therapist support ($\chi^2(3)$=15.69, P=.001). No differences in non-usage attrition were found among the 3 Featback conditions ($\chi^2(2)$=5.24, P=.07).

Baseline characteristics of participants are summarized in Table 1. Significant differences between the conditions were found regarding age, duration of ED psychopathology, and marital status, whereas no significant differences were found for any other baseline variables. No significant differences between the study conditions were found regarding the number of psychological health care appointments received (ie, appointments with a dietitian, social worker, psychologist, psychiatrist, or psychotherapist) during the intervention period (F(3,245)=0.29, P=.84). One hundred participants (40.2%) did not receive any psychological health care appointments during this period, whereas 149 participants (59.8%) did have such appointments (range 1-40).

Table 1. Baseline characteristics (nonimputed) of the study population; data are provided as means (SD) or numbers (percentages)

Characteristics	Featback (n=87)	Featback+ low-intensity therapist support (n=88)	Featback+ high-intensity therapist support (n=89)	Waiting list control condition (n=90)	Total sample (n=354)	Statistics
Gender						$\chi^2(3)=2.02, P=.57$
Male	1 (1.1%)	1 (1.1%)	2 (2.2%)	0 (0.0%)	4 (1.1%)	
Female	86 (98.9%)	87 (98.9%)	87 (97.8%)	90 (100.0%)	350 (98.9%)	
Marital status						$\chi^2(6)=13.22, P=.04$
Married or living together	28 (32.2%)[a,b]	17 (19.3%)[a,b]	21 (23.6%)[b]	11 (12.2%)[a]	77 (21.8%)	
Single or living alone	58 (66.7%)	71 (80.7%)	67 (75.3%)	79 (78.8%)	275 (77.7%)	
Divorced	1 (1.1%)	0 (0.0%)	1 (1.1%)	0 (0.0%)	2 (0.6%)	
Education level						$\chi^2(6)=7.69, P=.26$
Low	4 (4.6%)	4 (4.5%)	7 (7.9%)	10 (11.1%)	25 (7.1%)	
Intermediate	16 (18.4%)	26 (29.5%)	19 (21.3%)	17 (18.9%)	78 (22.0%)	
High	67 (77.0%)	58 (65.9%)	63 (70.8%)	63 (70.0%)	251 (70.9%)	
Use of psychotropic medication						$\chi^2(3)=3.35, P=.34$
Yes	21 (24.7%)	17 (19.5%)	16 (18.2%)	25 (28.4%)	79 (22.7%)	
No	64 (75.3%)	70 (80.5%)	72 (81.8%)	63 (71.6%)	269 (77.3%)	
Employment status						$\chi^2(9)=8.96, P=.44$
School or study	50 (58.1%)	48 (55.2%)	40 (45.5%)	51 (56.7%)	189 (53.8%)	
Employed	25 (29.1%)	22 (25.3%)	35 (39.8%)	30 (33.3%)	112 (31.9%)	
Unemployed or homemaker	4 (4.7%)	8 (9.2%)	4 (4.5%)	3 (3.3%)	19 (5.4%)	
Sick leave or disabled	7 (8.1%)	9 (10.3%)	9 (10.2%)	6 (6.7%)	31 (8.8%)	
Treatment history for ED[c]						$\chi^2(3)=4.43, P=.22$
Yes	48 (55.2%)	40 (45.5%)	39 (43.8%)	36 (40.0%)	163 (46.0%)	
No	39 (44.8%)	48 (54.5%)	50 (56.2%)	54 (60.0%)	191 (54.0%)	
Age (years)	24.7 (7.1)[a,b]	23.0 (7.0)[a]	26.3 (9.2)[b]	22.8 (6.6)[a]	24.2 (7.7)	$F(3,350)=4.17, P=.01$
Body mass index	21.8 (5.0)	21.2 (4.8)	21.4 (5.4)	20.6 (4.6)	21.2 (5.0)	$F(3,347)=1.03, P=.38$
Duration of ED problems (years)	8.1 (6.9)[a,b]	6.5 (5.8)[a,b]	8.2 (7.7)[b]	5.7 (5.6)[a]	7.1 (6.6)	$F(3,346)=3.05, P=.03$
Global ED psychopathology (EDE-Q)[d]	4.2 (0.8)	4.4 (0.9)	4.0 (0.8)	4.1 (1.1)	4.2 (0.9)	$F(3,113)=1.54, P=.21$
AN[e] psychopathology (SEED[f]-AN)	1.1 (0.4)	1.1 (0.4)	1.1 (0.4)	1.1 (0.4)	1.1 (0.4)	$F(3,347)=0.24, P=.87$
BN[g] psychopathology (SEED-BN)	1.4 (0.7)	1.5 (0.7)	1.5 (0.6)	1.5 (0.7)	1.5 (0.7)	$F(3,349)=0.30, P=.82$

[a,b]Significant group differences were further investigated using Bonferroni post-hoc comparisons; different superscript letters indicate significant differences between the conditions.

[c]ED: eating disorder.

[d]EDE-Q: Eating Disorder Examination Questionnaire.

[e]AN: anorexia nervosa.

[f]SEED: Short Evaluation of Eating Disorders.

[g]BN: bulimia nervosa.

Participants in this study demonstrated severe levels of ED psychopathology: their EDE-Q scores were comparable to the overall norm for treatment-seeking patients with an ED in our clinical program [40]. The mean EDE-Q score of 4.2 (SD=0.9) is furthermore markedly above the clinical threshold, as recent literature demonstrated reliable EDE-Q cutoff scores of >2.50

[41] and >2.12 [42]. Approximately 98.6% (n=349) of the study participants scored above the cutoff score of 2.5. To provide a diagnostic impression of the study sample, we used the EDE-Q to approximate diagnostic classifications according to the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) [43]. Subsequently, 103 (29%) participants demonstrated symptoms of anorexia nervosa, being a body mass index of ≤18.5 combined with a fear of weight gain or of becoming fat. A total of 93 participants (26%) reported binge eating disorder symptoms: binge eating episodes once a week or more during the past 28 days, without recurrent use of inappropriate compensatory behaviors (ie, less than once a week over the past 28 days). Seventy-seven (22%) participants reported symptoms of bulimia nervosa, being episodes of binge eating and inappropriate compensatory behaviors both at least once a week or more during the past 28 days. Only 14 participants (4%) demonstrated symptoms of purging disorder, that is, purging behaviors once a week or more during the past 28 days in the absence of binge eating episodes. Finally, 5 participants (14%) reported ED symptoms that may be classified as "unspecified feeding or ED," or ED problems without a DSM-5 classification. Seventeen participants (5%) could not be classified owing to missing data regarding binge eating episodes or body mass index. The 4 study conditions did not differ with respect to the type of ED (χ^2(15)=19.33, *P*=.20).

Intervention Compliance

Participants in the 3 Featback conditions completed a mean number of 5.6 (SD=2.3, range 0-8) of 8 weekly monitoring questionnaires, with no significant difference between the conditions (F(2,261)=1.36, *P*=.258). Participants in the 2 Featback conditions with therapist support received a total of 1407 support sessions, with email being the most popular medium (n=937, 67%), followed by chat (n=417, 30%) and teleconference (n=53, 4%). These proportions of email (*t* (1,155)=−1.63, *P*=.11), chat (*t* (1,153)=1.42, *P*=.16), and teleconference (*t* (1,159)=0.53, *P*=.59) were similar for the 2 study conditions. The mean number of received therapist support sessions differed significantly between Featback with low- and high-intensity therapist support (*t* (175)=8.24, *P*<.001): participants in the former condition received on average 4.7

(SD=2.7, range 0-8) sessions, whereas participants in the latter condition received on average 11.2 (SD=6.9, range 0-24) sessions. Thus, we successfully created 2 different intervention conditions regarding the intensity of therapist support.

Comparison of Intervention Conditions With Waiting List Condition

The outcome data for each of the 4 conditions over time can be found in Multimedia Appendix 2. Table 2 summarizes the results of the mixed model analyses comparing the 3 Featback conditions with the WLC (statistical model 1). As summarized in Table 2, from baseline to post-intervention, significant time-by-condition effects were found for bulimic psychopathology (d=−0.16, 95% CI=−0.31 to −0.01), symptoms of depression and anxiety (d=−0.31, 95% CI=−0.54 to −0.09), and perseverative thinking (d=−0.28, 95% CI=−0.45 to −0.11). These interaction effects indicated greater reductions in psychopathology for participants in the Featback conditions as compared with the WLC. For global ED psychopathology and ED-related quality of life, only significant time effects were found, indicating improvements over time. From post-intervention to 3-month follow-up, significant time-by-condition effects were found for ED-related quality of life (d=−0.22, 95% CI=−0.38 to −0.06) and symptoms of depression and anxiety (d=−0.21, 95% CI=−0.33 to −0.09), indicating more improvements in the Featback conditions as compared with the WLC during the 3-month follow-up period (see Table 2). For anorectic and bulimic psychopathology and levels of perseverative thinking, no interaction effects were found, but significant time effects were found that indicated improvements over time. Completer analyses confirmed the conclusions of the intent-to-treat analyses and are therefore not reported.

Comparison of Active Intervention Conditions

In statistical models 2 and 3, we compared the intervention conditions and thus investigated the added value of therapist support, and higher versus lower intensities of therapist support, respectively. As shown in Multimedia Appendices 3 and 4, participants in all Featback conditions

XSL•FO
RenderX

Figure 1. CONSORT diagram: Flow of participants through each stage of the randomized controlled trial.

XSL•FO
RenderX
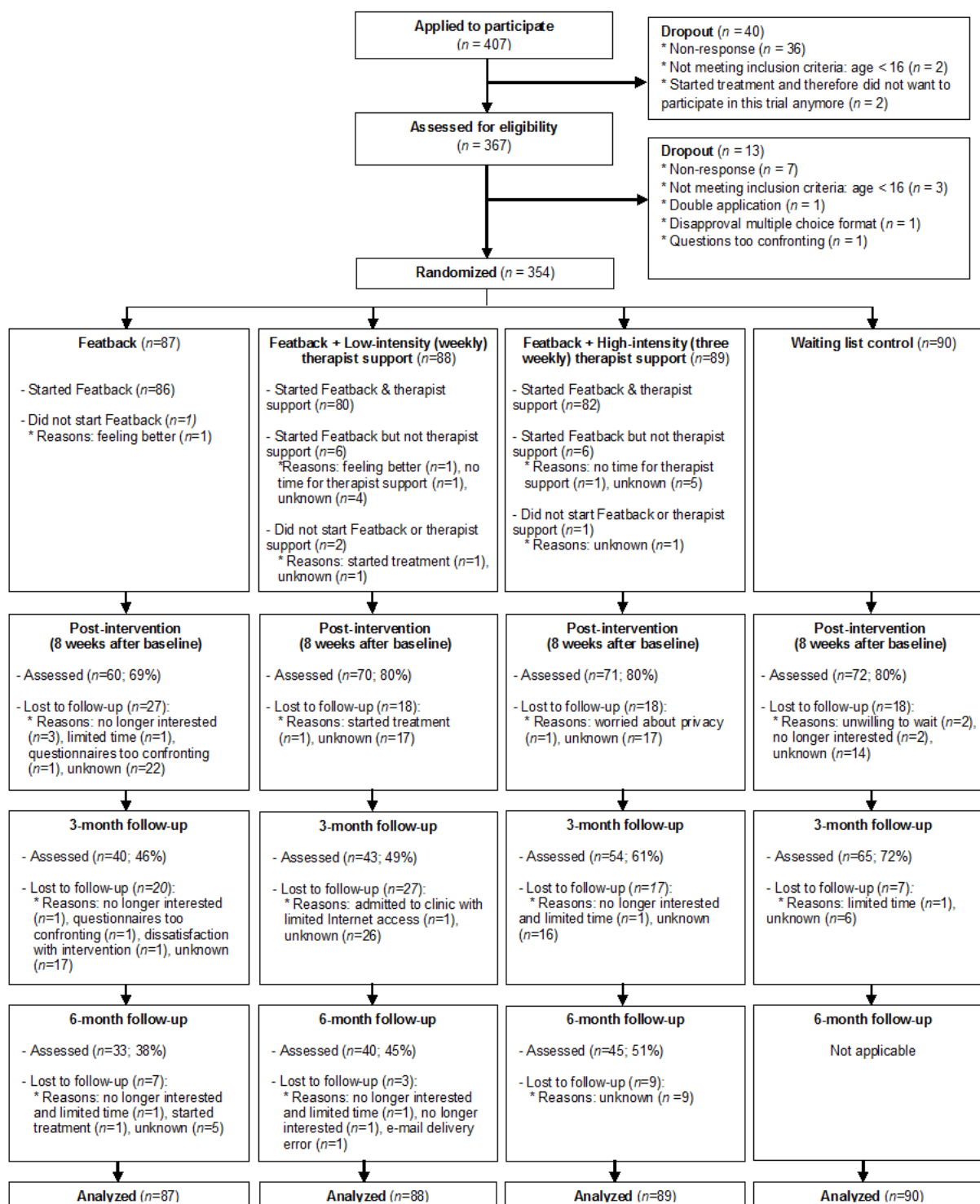
Table 2. Results of linear mixed model analyses comparing the effectiveness of an Internet-based fully automated monitoring and feedback intervention with a waiting list control condition; results are based on the pooled results of 100 multiple imputed datasets.

Measure	Time effects			Time × condition effects		
	B	$t (P)$	95% CI	B	$t (P)$	95% CI
Anorectic psychopathology (SEED[a]-AN[b])						
Baseline to post-intervention	−0.02	−0.42 (.44)	−0.06 to 0.03	0.01	0.35 (.73)	−0.04 to 0.06
Post-intervention to 3-month follow-up	−0.05	−2.21 (.03)	−0.10 to −0.006	0.04	1.41 (.16)	−0.02 to 0.09
Bulimic psychopathology (SEED-BN[c])						
Baseline to post-intervention	−0.07	−1.50 (.11)	−0.15 to 0.02	−0.11	−2.13 (.03)	−0.21 to −0.009
Post-intervention to 3-month follow-up	−0.12	−2.51 (.01)	−0.22 to −0.03	−0.02	−0.42 (.67)	−0.14 to 0.09
Global ED psychopathology (EDE-Q[d])						
Baseline to post-intervention	−0.22	−3.07 (.002)	−0.37 to −0.08	−0.09	−1.08 (.28)	−0.26 to 0.08
Post-intervention to 3-month follow-up	−0.18	−2.44 (.02)	−0.32 to −0.03	−0.07	−0.77 (.44)	−0.25 to 0.11
ED-related quality of life (ED-QOL[e])						
Baseline to post-intervention	−0.13	−3.46 (.001)	−0.20 to −0.06	−0.03	−0.74 (.46)	−0.12 to 0.05
Post-intervention to 3-month follow-up	−0.06	−1.44 (.15)	−0.14 to 0.02	−0.13	−2.70 (.007)	−0.23 to −0.04
Symptoms anxiety & depression (PHQ-4[f])						
Baseline to post-intervention	−0.37	−1.92 (.06)	−0.74 to 0.007	−0.94	−4.11 (<.001)	−1.39 to −0.49
Post-intervention to 3-month follow-up	−0.29	−1.43 (.15)	−0.69 to 0.11	−0.62	−2.53 (.01)	−1.11 to −0.14
Perseverative thinking (PTQ[g])						
Baseline to post-intervention	−0.08	−1.48 (.14)	−0.18 to 0.03	−0.20	−3.20 (.001)	−0.32 to −0.07
Post-intervention to 3-month follow-up	−0.16	−2.89 (.004)	−0.26 to −0.05	−0.05	−0.82 (.41)	−0.18 to 0.07

[a]SEED: Short Examination of Eating Disorders.

[b]AN: anorexia nervosa.

[c]BN: bulimia nervosa.

[d]EDE-Q: Eating Disorder Examination Questionnaire.

[e]ED-QOL: Eating Disorder–related Quality Of Life.

[f]PHQ-4: 4-item Patient Health Questionnaire.

[g]PTQ: Perseverative Thinking Questionnaire.

improved over time (baseline vs post-intervention, and post-intervention vs 3- and 6-month follow-up, respectively) with respect to bulimic psychopathology, global ED psychopathology, ED-related quality of life, symptoms of depression and anxiety, and levels of perseverative thinking (all P values ≤.01). When comparing Featback without therapist support with the pooled Featback conditions with therapist support (statistical model 2), the results demonstrated no significant differences between the conditions over time (all P values >.05, see Multimedia Appendix 3), indicating that participants improved to a similar degree. When comparing Featback with low- versus high-intensity therapist support (statistical model 3), no significant time-by-condition effects were found for most of the outcome measures (Multimedia Appendix 4). Except for ED-related quality of life, participants who received Featback with high-intensity therapist support showed greater improvements in ED-related quality of life from baseline to post-intervention (P=.001, d=0.15, 95% CI=0.06-0.24) and from post-intervention to 6-month follow-up (P=.01, d=0.14, 95% CI 0.03-0.25) than participants who

received Featback with low-intensity therapist support. This finding should be interpreted with caution because participants who received Featback without therapist support scored in between and thereby not significantly different from the 2 Featback conditions with therapist support (Multimedia Appendix 2). Completer analyses confirmed the conclusions of the intent-to-treat analyses and are therefore not reported.

Participants' Experiences

Regarding participants' experiences, significant differences in participants' level of satisfaction with Featback were found (F(2,184)=38.41, P<.001). Participants who received Featback without therapist support were significantly less satisfied (M=5.0, SD=1.9, scale 1-10) than participants who received Featback with low- (M=7.1, SD=1.5) or high-intensity therapist support (M=7.4, SD=1.3), whereas no differences between the latter 2 were found. Overall, participants were very satisfied with the therapist support (M=8.0, SD=1.4, scale 1-10), with no significant differences between the low- and high-intensity therapist support conditions (t (1,117)= −0.34, P=.74). In

XSL•FO

RenderX

addition, no significant differences in satisfaction with the different therapists were found (F(6,112)=0.36, *P*=.902).

A total of 158 participants provided negative feedback, and 160 participants provided positive feedback to the open-ended questions regarding their satisfaction with the intervention. Participants' most reported critical comments included statements about the limitations of the automated feedback (n=95, 60%), for example, it being too general or impersonal, as well as the lack of more personal or individual therapist support. Most of the positive comments (n=107, 84.3%) included complementary remarks regarding the individual therapist support, such as participants having received good advice and support, having enjoyed the empathy, warmth, and attention of the therapists, as well as the feeling that someone was looking after them. Approximately one third (n=45, 28%) of all positive comments included positive feedback on this system, for example, experiencing the system as a good checkup supporting moments of reflection. No adverse effects from Featback were reported.

Discussion

To our knowledge, this is the first randomized controlled trial to investigate an Internet-based fully automated self-monitoring and feedback intervention (Featback) and the added value of 2 different intensities of therapist support for individuals with ED psychopathology. The results demonstrated Featback to be superior to a WLC in reducing bulimic psychopathology (ie, a total severity index of binge eating, compensatory behaviors, and overconcern with body shape and weight), perseverative thinking, and symptoms of depression and anxiety. Thus, self-monitoring of ED-related attitudes and behaviors and receiving feedback by means of an automatic system can be effective in reducing psychopathology. No effects were found regarding anorectic psychopathology; hence, Featback may be more suitable for individuals with bulimic psychopathology. Interestingly, when comparing Featback with and without therapist support, no added value was found for therapist support in terms of the effectiveness of the intervention, although participants who received Featback with therapist support were significantly more satisfied.

Our findings add to the growing body of literature indicating the potential of ehealth interventions for individuals with (ED) psychopathology [8,9,12,13,44]. Our results are furthermore in line with 2 studies demonstrating that interventions supplemented with automated support can be equally effective to human support [18,19]. A fully automated Internet-based intervention such as Featback is a promising, widely disseminable, easily accessible, and potentially effective means of providing care for individuals with ED psychopathology. Such care is particularly important for these individuals, given that many do not seek or receive appropriate mental health care [6]. Hence, Internet-based self-help interventions might help to bridge the gap between mental disorders and mental health care services, by improving the help-seeking pathways. Internet-based automated self-monitoring and feedback systems may be of interest to a number of *other* areas in the field of psychiatry. Indeed, a recent study [45] demonstrated an Internet-based intervention including self-monitoring via text messages to be effective in remitted patients with symptoms of depression.

The finding that Featback was equally effective with and without therapist support is in line with that of several previous studies [46-48], however, in contrast to the result of a recent meta-analysis that included Internet-based interventions for a range of mental health problems [16]. This meta-analysis demonstrated guided Internet-based interventions to be significantly superior to unguided interventions. However, the larger effect sizes in the guided interventions may have been biased by significantly higher adherence rates in the guided interventions as compared with unguided interventions [16], whereas adherence rates in our study were similar for the guided and unguided conditions. A possible explanation for why therapist support did not enhance the effectiveness of Featback is that the monitoring and feedback system alone was already a relatively powerful intervention in reducing ED symptoms. Self-monitoring is an important clinical technique that is often used in cognitive behavioral therapy [20]. It can help an individual to gain a more comprehensive understanding of one's psychopathology. By self-monitoring one's psychopathology and receiving feedback, an individual is stimulated to think about the frequency, antecedents, and consequences of their problematic behaviors and attitudes [20]. Furthermore, through the provided feedback, individuals are encouraged to think about possible solutions to achieve positive behavioral changes, and in addition, the feedback can help them in applying and developing certain skills to promote such behavioral changes in their daily lives. It could be speculated that the self-monitoring and feedback system of the Featback intervention already provided such a powerful intervention to help reduce ED psychopathology that the therapist support did not add an extra effect. Within this context, the individual therapist support might primarily be appreciated for its empathy, warmth, and attention, as well as the feeling that someone is looking after you and listening to you.

Increasing the frequency of therapist support did not significantly affect outcome, which is in line with the results of a study that experimentally investigated different intensities of therapist support in an Internet-based treatment for panic disorder [17]. More frequent therapist support did furthermore not affect the participants' satisfaction with the intervention or their therapist. Thus, increasing the amount of therapist contact may not necessarily result in increased effectiveness or increased satisfaction with Internet-based interventions. Nevertheless, future dose-response studies should replicate these rather unexpected findings before any firm conclusions can be drawn with respect to the added value of different intensities of therapist support. Also, cost-effectiveness studies comparing different intensities of therapist support would be of great interest. Such studies can facilitate decision making on how to most optimally deliver therapist support within Internet-based interventions. How much money needs to be invested in terms of additional therapist support to realize a particular amount of additional improvement in health outcomes? And does the extra benefit resulting from therapist support justify the extra cost:

is adding a certain amount of therapist support good value for money?

Interestingly, our results show a discrepancy between the added value of therapist support in terms of effectiveness (no added value of therapist support) and satisfaction with the intervention (added value of therapist support). The fact that therapist support did increase the satisfaction of participants significantly might well be due to the empathy, warmth, and attention of the therapists. Individuals with ED are often ashamed about their ED and can feel isolated and unsupported, as well as misunderstood by their personal environment [49]. Although the automated feedback as part of Featback expresses interest in participants' well-being and provides advice on how to possibly counteract certain dysfunctional beliefs or behaviors, it is not interactive. That is, individuals are not able to share their personal story, history, in-depth feelings and emotions, or experiences. In the individual therapist support sessions, they were able to (anonymously) ventilate their problems and emotions, and the majority reported on how nice it was to have someone looking after them, understanding them, and listening to them. Translating these study results to everyday clinical practice is challenging, given the added value of therapist regarding satisfaction, but not effectiveness. The resulting dilemma is about how to implement Featback: with or without therapist support? Adding such support implies more costs while not necessarily resulting in increased effectiveness. That being said, adding therapist support presumably heightens the attractiveness and thus reach of the intervention, eventually leaving more individuals feeling supported. An interesting future research direction would be to investigate the effectiveness of adding personal support by means of a Web-based peer support group. Possibly, the personal interactive support of peers might be sufficient to increase satisfaction rates, while at the same time reducing costs in comparison to trained professionals.

Adding therapist support did not enhance study adherence because no differences between study dropout rates were found between the 3 Featback conditions. However, our results showed that at 3-month follow-up, participants in the waiting list condition dropped out less often than participants who received Featback without or with low-intensity therapist support. Presumably, participants in the waiting list condition were more motivated to complete the study questionnaires given their knowledge that they would receive Featback with low-intensity therapist support after completing this follow-up questionnaire.

It is noteworthy that Featback produced significant reductions in psychopathology over and above usual care. Participants' treatment status (yes or no) or number of received psychological health care appointments during the intervention period did not significantly differ between the study conditions and could furthermore not account for the superiority of Featback in comparison to WLC when entered as a predictor in the model. This suggests minimal self-help interventions such as Featback to be of interest for a broad population of individuals with ED symptoms. The small effect sizes match our expectations, given the type of intervention (ie, self-help) and the fact that most participants received psychological health care during the intervention period. Interventions such as Featback could be incorporated within regular treatment settings (ie, blended care),

where it would enable accurate monitoring of patients' well-being in treatment settings and in their everyday lives [20,50]. Also, information about patterns of dysfunctional attitudes and behaviors as gathered by the use of self-monitoring may aid in clarifying the rationale and goals for treatment, as well as informing therapists and patients about the patient's progress in treatment. It could furthermore be useful to incorporate self-help interventions such as Featback as a first step within a stepped-care approach in the treatment of ED, thereby providing low-intensity care to individuals with ED symptoms who might not (yet) need more intense specialist care. Individuals who remain symptomatic after a certain period of time could then "step up" to a more intense specialist care. Similarly, Featback could also be used as a "step down" intervention after a more intensive treatment. Individuals can keep track of their ED symptoms and can be supported in their process of recovery. In addition, Featback as a "step down" intervention could allow for early identification and prevention of relapse. The potential effectiveness and cost-effectiveness of a stepped-care approach starting with self-help, as compared with cognitive behavior therapy, has already been demonstrated in a large multicenter trial for individuals with bulimia nervosa [51,52]. In sum, investigating the effectiveness of Featback within treatment settings, or as part of stepped-care approaches in the treatment of ED, is an interesting area for future research.

This study has several strengths and limitations. Strengths include the large sample size, randomized controlled design, intent-to-treat analyses, and the use of multiple imputation methods as these have shown improved performance over alternative approaches such as complete case analysis or single imputation methods [36]. Limitations include the lack of a 6-month follow-up for the WLC and the considerable amount of missing data at 3- and 6-month follow-up. The non-significant differences between the 3 Featback conditions should be interpreted with caution as statistical power might have been reduced due to the missing data at 3- and 6-month follow-up. The use of broad eligibility criteria can be regarded as both a strength and a limitation. The broad inclusion criteria may well have led to a study population that bears close resemblance to reality, thereby enhancing the generalizability of our findings and being consistent with the aim of an easily accessible intervention for a broad population of individuals with ED psychopathology. Alternatively, the broad inclusion criteria can be regarded as a limitation given the potential influences of variables such as the presence of comorbid disorders or the use of co-interventions on study outcome measures that were not under study control. Nevertheless, we attempted to reduce the risk of bias by acquiring detailed information on participant characteristics and external influences, so that these influences could be examined and controlled for in the analyses. Finally, the use of Web-based self-report assessments can be considered both a strength and limitation. Advantages include a reduction in research costs and being in line with the aims of the anonymous ehealth intervention: being able to remain anonymous, which lowers the barriers of seeking help and maximizing the accessibility, efficiency, and availability of health care services. Another advantage includes the minimization of the risk of bias because of the lack of face-to-face contact with participants. However, the latter might

XSL•FO
RenderX

have reduced study and/or intervention commitment [9], and it resulted in the absence of a face-to-face diagnostic interview. Although we did provide a diagnostic impression of the study sample using the EDE-Q [43], it must be emphasized that the resulting classifications provide only an approximation of DSM-5 classifications as there are limitations to the use of the EDE-Q in evaluating the diagnostic criteria of ED [53].

In conclusion, an Internet-based fully automated monitoring and feedback intervention was effective in reducing psychopathology and is an interesting means of providing care for individuals with ED symptoms. Supplemental therapist support enhanced satisfaction with the intervention but did not increase its effectiveness. An interesting next step is to economically evaluate Featback with and without therapist support to determine its cost-effectiveness in comparison to a waiting list. Also, examining potential predictors, moderators, and mediators of intervention response will help to inform the field regarding for whom and how Featback work(s). A final topic for future investigation is a focus on opening the black box of therapeutic support in Internet-based interventions: what do therapists actually do when providing Web-based support and can their behavior be linked to the effectiveness of such interventions?

Acknowledgments

Conflicts of Interest

None declared.

Multimedia Appendix 1

Specification of statistical models and contrast coding.

[PDF File (Adobe PDF File), 17KB - jmir_v18i6e159_app1.pdf]

Multimedia Appendix 2

Non-imputed outcome data (means and standard deviations) in a trial investigating the effectiveness of Internet-based fully automated monitoring and feedback intervention "Featback" with different intensities of therapist support and a waiting list control condition.

[PDF File (Adobe PDF File), 20KB - jmir_v18i6e159_app2.pdf]

Multimedia Appendix 3

Results of linear mixed model analyses comparing the effectiveness of an Internet-based fully automated monitoring and feedback intervention with and without therapist support (statistical model 2). Results are based on the pooled results of 100 multiple imputed datasets.

[PDF File (Adobe PDF File), 21KB - jmir_v18i6e159_app3.pdf]

Multimedia Appendix 4

Results of linear mixed model analyses comparing the effectiveness of an Internet-based fully automated monitoring and feedback intervention with low-intensity (once a week) versus high-intensity (3 times a week) therapist support (statistical model 3). Results are based on the pooled results of 100 multiple imputed datasets.

[PDF File (Adobe PDF File), 22KB - jmir_v18i6e159_app4.pdf]

References

1. Berkman ND, Lohr KN, Bulik CM. Outcomes of eating disorders: a systematic review of the literature. Int J Eat Disord 2007 May;40(4):293-309. [doi: 10.1002/eat.20369] [Medline: 17370291]
2. Hudson JI, Hiripi E, Pope HG, Kessler RC. The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. Biol Psychiatry 2007 Feb 1;61(3):348-358 [FREE Full text] [doi: 10.1016/j.biopsych.2006.03.040] [Medline: 16815322]

3. Keel PK, Brown TA. Update on course and outcome in eating disorders. Int J Eat Disord 2010 Apr;43(3):195-204. [doi: 10.1002/eat.20810] [Medline: 20186717]

4. Smink FR, van Hoeken D, Hoek HW. Epidemiology, course, and outcome of eating disorders. Curr Opin Psychiatry 2013 Nov;26(6):543-548. [doi: 10.1097/YCO.0b013e328365a24f] [Medline: 24060914]

5. Kessler RC, Berglund PA, Chiu WT, Deitz AC, Hudson JI, Shahly V, et al. The prevalence and correlates of binge eating disorder in the World Health Organization World Mental Health Surveys. Biol Psychiatry 2013 May 1;73(9):904-914 [FREE Full text] [doi: 10.1016/j.biopsych.2012.11.020] [Medline: 23290497]

6. Hart LM, Granillo M, Jorm A, Paxton S. Unmet need for treatment in the eating disorders: a systematic review of eating disorder specific treatment seeking among community cases. Clin Psychol Rev 2011 Jul;31(5):727-735. [doi: 10.1016/j.cpr.2011.03.004] [Medline: 21501580]

7. Becker AE, Hadley AA, Perloe A, Fay K, Striegel-Moore RH. A qualitative study of perceived social barriers to care for eating disorders: perspectives from ethnically diverse health care consumers. Int J Eat Disord 2010 Nov 1;43(7):633-647 [FREE Full text] [doi: 10.1002/eat.20755] [Medline: 19806607]

8. Beintner I, Jacobi C, Taylor CB. Effects of an Internet-based prevention programme for eating disorders in the USA and Germany--a meta-analytic review. Eur Eat Disord Rev 2012 Jan;20(1):1-8. [doi: 10.1002/erv.1130] [Medline: 21796737]

9. Aardoom JJ, Dingemans AE, Spinhoven P, Van Furth EF. Treating eating disorders over the internet: a systematic review and future research directions. Int J Eat Disord 2013 Sep;46(6):539-552. [doi: 10.1002/eat.22135] [Medline: 23674367]

10. Melioli T, Bauer S, Franko DL, Moessner M, Ozer F, Chabrol H, et al. Reducing eating disorder symptoms and risk factors using the internet: A meta-analytic review. Int J Eat Disord 2016 Jan;49(1):19-31. [doi: 10.1002/eat.22477] [Medline: 26607683]

11. Aardoom JJ, Dingemans AE, Van Furth EF. E-Health Interventions for Eating Disorders: Emerging Findings, Issues, and Opportunities. Curr Psychiatry Rep 2016 Apr;18(4):42. [doi: 10.1007/s11920-016-0673-6] [Medline: 26946513]

12. Bauer S, Papezova H, Chereches R, Caselli G, McLoughlin O, Szumska I, et al. Advances in the prevention and early intervention of eating disorders: The potential of Internet-delivered approaches. Mental Health & Prevention 2013 Dec;1(1):26-32. [doi: 10.1016/j.mhp.2013.10.003]

13. Loucas CE, Fairburn CG, Whittington C, Pennant ME, Stockton S, Kendall T. E-therapy in the treatment and prevention of eating disorders: A systematic review and meta-analysis. Behav Res Ther 2014 Oct 5;63C:122-131 [FREE Full text] [doi: 10.1016/j.brat.2014.09.011] [Medline: 25461787]

14. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]

15. Spek V, Cuijpers P, Nyklícek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. Psychol Med 2007 Mar;37(3):319-328. [doi: 10.1017/S0033291706008944] [Medline: 17112400]

16. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on Internet-based mental health interventions: A systematic review. Internet Interventions 2015 Oct;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003]

17. Klein B, Austin D, Pier C, Kiropoulos L, Shandley K, Mitchell J, et al. Internet-based treatment for panic disorder: does frequency of therapist contact make a difference? Cogn Behav Ther 2009;38(2):100-113. [doi: 10.1080/16506070802561132] [Medline: 19306149]

18. 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. [doi: 10.1001/archinte.166.15.1620] [Medline: 16908795]

19. Kelders SM, Bohlmeijer E, Pots WT, van Gemert-Pijnen JE. Comparing human and automated support for depression: Fractional factorial randomized controlled trial. Behav Res Ther 2015 Sep;72:72-80. [doi: 10.1016/j.brat.2015.06.014] [Medline: 26196078]

20. Cohen JS, Edmunds JM, Brodman DM, Benjamin CL, Kendall PC. Using self-monitoring: Implementation of collaborative empiricism in cognitive-behavioral therapy. Cognitive and Behavioral Practice 2013 Nov;20(4):419-428. [doi: 10.1016/j.cbpra.2012.06.002]

21. Fairburn CG, Cooper Z, Shafran R. Cognitive behaviour therapy for eating disorders: a "transdiagnostic" theory and treatment. Behav Res Ther 2003 May;41(5):509-528. [Medline: 12711261]

22. Aardoom JJ, Dingemans AE, Spinhoven P, Hakkaart-van Roijen L, Van Furth EF. An Internet-based intervention for eating disorders consisting of automated computer-tailored feedback with or without supplemented frequent or infrequent support from a coach: study protocol for a randomized controlled trial. Trials 2013;14:340 [FREE Full text] [doi: 10.1186/1745-6215-14-340] [Medline: 24135131]

23. Featback. URL: http://www.featback.nl/ [accessed 2016-03-01] [WebCite Cache ID 6fgI6ha4A]

24. Proud2Bme. URL: http://www.proud2bme.nl [accessed 2016-03-01] [WebCite Cache ID 6fgHnixVa]

25. Killen JD, Taylor CB, Hayward C, Wilson DM, Haydel KF, Hammer LD, et al. Pursuit of thinness and onset of eating disorder symptoms in a community sample of adolescent girls: a three-year prospective analysis. Int J Eat Disord 1994 Nov;16(3):227-238. [Medline: 7833956]

XSL·FO

RenderX

26. Bauer S, Winn S, Schmidt U, Kordy H. Construction, scoring and validation of the Short Evaluation of Eating Disorders (SEED). Eur Eat Disord Rev 2005 May;13(3):191-200. [doi: 10.1002/erv.637]

27. Bauer S, Moessner M, Wolf M, Haug S, Kordy H. ES[S]PRIT - an Internet-based programme for the prevention and early intervention of eating disorders in college students. Br J Guid Couns 2009 Aug;37(3):327-336. [doi: 10.1037/a0025806]

28. Schlaken F, Wolters W, Tilanus M, van Gemert M, Hoogenhuyze C, Meijer E, et al. Handbook of Online Counseling [Handboek online hulpverlening]. Houten: Bohn Stafleu van Loghum; 2010.

29. Fairburn CG, Beglin S. Eating Disorder Examination Questionnaire (EDE-Q 6.0). In: Fairburn CG, editor. Cognitive behavior therapy and eating disorders. New York: Guilford Press; 2008:309-313.

30. Berg KC, Peterson CB, Frazier P, Crow SJ. Psychometric evaluation of the eating disorder examination and eating disorder examination-questionnaire: A systematic review of the literature. Int J Eat Disord 2012 Apr;45(3):428-438. [Medline: 21744375] [doi: 10.1002/eat.20931]

31. Engel SG, Wittrock DA, Crosby RD, Wonderlich SA, Mitchell JE, Kolotkin RL. Development and psychometric validation of an eating disorder-specific health-related quality of life instrument. Int J Eat Disord 2006 Jan;39(1):62-71. [doi: 10.1002/eat.20200] [Medline: 16345055]

32. Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. Psychosomatics 2009;50(6):613-621. [doi: 10.1176/appi.psy.50.6.613] [Medline: 19996233]

33. Ehring T, Raes F, Weidacker K, Emmelkamp PMG. Validation of the Dutch Version of the Perseverative Thinking Questionnaire (PTQ-NL). European Journal of Psychological Assessment 2012;28:102-108. [doi: 10.1027/1015-5759/a000097]

34. Hakkaart-van Roijen L, Donker M, Tiemens B. Handleiding trimbos/iMTA questionnaire for costs associated with psychiatric illness (TiC-P). Rotterdam: Institute for Medical Technology Assessment; 2002.

35. Doove LL, Van Buuren S, Dusseldorp E. Recursive partitioning for missing data imputation in the presence of interaction effects. Computational Statistics and Data Analysis 2014;72:92-104. [doi: 10.1016/j.csda.2013.10.025]

36. Schafer JL, Graham JW. Missing data: our view of the state of the art. Psychol Methods 2002 Jun;7(2):147-177. [Medline: 12090408] [doi: 10.1037/1082-989X.7.2.147]

37. Stevens J. Applied multivariate statistics for the social sciences. Hillsdale, NJ: Taylor & Francis; 2009.

38. Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: Wiley; 1987.

39. Feingold A. Confidence interval estimation for standardized effect sizes in multilevel and latent growth modeling. J Consult Clin Psychol 2015 Feb;83(1):157-168 [FREE Full text] [doi: 10.1037/a0037721] [Medline: 25181028]

40. Aardoom JJ, Dingemans AE, Slof Op't Landt MC, Van Furth EF. Norms and discriminative validity of the Eating Disorder Examination Questionnaire (EDE-Q). Eat Behav 2012 Dec;13(4):305-309. [doi: 10.1016/j.eatbeh.2012.09.002] [Medline: 23121779]

41. Rø Ø, Reas DL, Stedal K. Eating Disorder Examination Questionnaire (EDE-Q) in Norwegian Adults: Discrimination between Female Controls and Eating Disorder Patients. Eur Eat Disord Rev 2015 Sep;23(5):408-412. [doi: 10.1002/erv.2372] [Medline: 26094887]

42. Machado PP, Martins C, Vaz AR, Conceição E, Bastos AP, Gonçalves S. Eating disorder examination questionnaire: psychometric properties and norms for the Portuguese population. Eur Eat Disord Rev 2014 Nov;22(6):448-453. [doi: 10.1002/erv.2318] [Medline: 25175299]

43. American Psychiatric Association. Diagnosticstatistical manual of mental disorders (5th ed). Washington, DC: American Psychiatric Association; 2013.

44. Cuijpers P, van Straten A, Andersson G. Internet-administered cognitive behavior therapy for health problems: a systematic review. J Behav Med 2008 Apr;31(2):169-177 [FREE Full text] [doi: 10.1007/s10865-007-9144-1] [Medline: 18165893]

45. Kok G, Burger H, Riper H, Cuijpers P, Dekker J, van Marwijk H, et al. The Three-Month Effect of Mobile Internet-Based Cognitive Therapy on the Course of Depressive Symptoms in Remitted Recurrently Depressed Patients: Results of a Randomized Controlled Trial. Psychother Psychosom 2015 Feb 21;84(2):90-99. [doi: 10.1159/000369469] [Medline: 25721915]

46. Berger T, Caspar F, Richardson R, Kneubühler B, Sutter D, Andersson G. Internet-based treatment of social phobia: A randomized controlled trial comparing unguided with two types of guided self-help. Behav Res Ther 2011 Mar;49(3):158-169. [Medline: 21255767] [doi: 10.1016/j.brat.2010.12.007]

47. Rheker J, Andersson G, Weise C. The role of "on demand" therapist guidance vs no support in the treatment of tinnitus via the internet: A randomized controlled trial. Internet Interventions 2015 May;2(2):189-199. [doi: 10.1016/j.invent.2015.03.007]

48. Mohr DC, Duffecy J, Ho J, Kwasny M, Cai X, Burns MN, et al. A randomized controlled trial evaluating a manualized TeleCoaching protocol for improving adherence to a web-based intervention for the treatment of depression. PLoS One 2013;8(8):e70086 [FREE Full text] [doi: 10.1371/journal.pone.0070086] [Medline: 23990896]

49. Linville D, Brown T, Sturm K, McDougal T. Eating disorders and social support: perspectives of recovered individuals. Eat Disord 2012;20(3):216-231. [doi: 10.1080/10640266.2012.668480] [Medline: 22519898]

50. Tregarthen JP, Lock J, Darcy AM. Development of a smartphone application for eating disorder self-monitoring. Int J Eat Disord 2015 Nov;48(7):972-982. [doi: 10.1002/eat.22386] [Medline: 26213130]

XSL•FO

RenderX

51. Crow SJ, Agras WS, Halmi KA, Fairburn CG, Mitchell JE, Nyman JA. A cost effectiveness analysis of stepped care treatment for bulimia nervosa. Int J Eat Disord 2013 May;46(4):302-307 [FREE Full text] [doi: 10.1002/eat.22087] [Medline: 23354913]

52. Mitchell JE, Agras S, Crow S, Halmi K, Fairburn CG, Bryson S, et al. Stepped care and cognitive-behavioural therapy for bulimia nervosa: randomised trial. Br J Psychiatry 2011 May;198(5):391-397 [FREE Full text] [doi: 10.1192/bjp.bp.110.082172] [Medline: 21415046]

53. Berg KC, Stiles-Shields EC, Swanson SA, Peterson CB, Lebow J, Le GD. Diagnostic concordance of the interview and questionnaire versions of the eating disorder examination. Int J Eat Disord 2012 Nov;45(7):850-855 [FREE Full text] [doi: 10.1002/eat.20948] [Medline: 21826696]

Abbreviations

ED: eating disorder
EDE-Q: Eating Disorder Examination Questionnaire
ED-QOL: ED-Related Quality of Life Questionnaire
PHQ-4: 4-item Patient Health Questionnaire
PTQ: Perseverative Thinking Questionnaire
SD: standard deviation
SEED: Short Evaluation of Eating Disorders
WLC: waiting list control

XSL•FO
RenderX

Original Paper

Which Combinations of Techniques and Modes of Delivery in Internet-Based Interventions Effectively Change Health Behavior? A Meta-Analysis

Lenneke van Genugten[1,2], PhD; Elise Dusseldorp[1,3], PhD; Thomas Llewelyn Webb[4], PhD; Pepijn van Empelen[1], PhD

[1]Expertise Group Life Style, TNO, Leiden, Netherlands

[2]Department of Public Health, Erasmus Medical Center, Rotterdam, Netherlands

[3]Institute of Psychology, Leiden University, Leiden, Netherlands

[4]Department of Psychology, University of Sheffield, Sheffield, United Kingdom

Corresponding Author:
Elise Dusseldorp, PhD
Institute of Psychology
Leiden University
PO Box 9555
Leiden, 2300 RB
Netherlands
Phone: 31 715278046
Fax: 31 715278046
Email: elise.dusseldorp@fsw.leidenuniv.nl

Abstract

Background: Many online interventions designed to promote health behaviors combine multiple behavior change techniques (BCTs), adopt different modes of delivery (MoD) (eg, text messages), and range in how usable they are. Research is therefore needed to examine the impact of these features on the effectiveness of online interventions.

Objective: This study applies Classification and Regression Trees (CART) analysis to meta-analytic data, in order to identify synergistic effects of BCTs, MoDs, and usability factors.

Methods: We analyzed data from Webb et al. This review included effect sizes from 52 online interventions targeting a variety of health behaviors and coded the use of 40 BCTs and 11 MoDs. Our research also developed a taxonomy for coding the usability of interventions. Meta-CART analyses were performed using the BCTs and MoDs as predictors and using treatment success (ie, effect size) as the outcome.

Results: Factors related to usability of the interventions influenced their efficacy. Specifically, subgroup analyses indicated that more efficient interventions (interventions that take little time to understand and use) are more likely to be effective than less efficient interventions. Meta-CART identified one synergistic effect: Interventions that included barrier identification/ problem solving and provided rewards for behavior change reported an average effect size that was smaller (=0.23, 95% CI 0.08-0.44) than interventions that used other combinations of techniques (=0.43, 95% CI 0.27-0.59). No synergistic effects were found for MoDs or for MoDs combined with BCTs.

Conclusions: Interventions that take little time to understand and use were more effective than those that require more time. Few specific combinations of BCTs that contribute to the effectiveness of online interventions were found. Furthermore, no synergistic effects between BCTs and MoDs were found, even though MoDs had strong effects when analyzed univariately in the original study.

KEYWORDS

XSL•FO
RenderX

Introduction

Online interventions hold great promise for the promotion of health behavior. The Internet is used by many individuals to find health-related information [1]. Three potential advantages of online interventions are high reach, low costs, and convenience for users (eg, timely delivery) [2]. Various meta-analyses have shown that online interventions designed to promote health behavior change can be effective, but that the effectiveness of interventions varies considerably [3-6]. One source of variability is differences in the behavior change techniques (BCTs) that are used by interventions. Research points to the importance of using standard definitions of BCTs [7,8] and has started to identify which BCTs are effective and which are less so. Yet relatively few studies have sought to identify the effectiveness of BCTs in online interventions. One exception is Webb et al [3], who examined the effectiveness of online interventions using a taxonomy of BCTs adapted from Abraham and Michie [9]. Webb et al [3] found, based on univariate analyses, that several BCTs were associated with larger than average effect sizes. Specifically, stress management and general communication skills training had the strongest positive effects, while emotion control training and providing information about others' approval were not effective.

In addition to deciding which BCTs to use in an intervention, a second challenge for online interventions is how to attract users, encourage them to engage in the intervention and explore the website, and have them return for follow-up visits as necessary [10-12]. This process may be more complicated than traditional intervention methods (eg, a letter, flyer, or video), and it is likely that the usability—or user friendliness—of the intervention [13] has a substantial bearing on the efficacy of that intervention. Usability refers to how easily the features in the intervention are to use and how pleasant it is for the user to engage with the intervention [14]. However, it is unclear which factors influence how usable an intervention is. Therefore, the present meta-analysis aimed to identify factors that influence the usability of interventions, as well as how these factors are related to the effectiveness of interventions.

Interactions Between Intervention Factors

Webb et al also found that interventions were more effective when more BCTs were included (see also [15]), suggesting that combining BCTs may be more effective than using one or two BCTs in isolation. Indeed, evidence suggests that BCTs can interact and have cumulative (or potentially synergistic) effects. For example, the combination of fear arousal and providing skill information has been shown to be particularly effective in promoting a variety of health behaviors, such as smoking or vaccination [16,17]. Similarly, Michie et al [18] found that interventions that combined self-monitoring with at least one other BCT specified by control theory (eg, goal setting) tended to have larger effects than interventions that used self-monitoring in isolation.

Dusseldorp et al [19] developed and applied a new method for looking at the effectiveness of combinations of moderators (eg, BCTs), which they called meta-CART [20]. Meta-CART combines "Classification And Regression Trees" (CART) and

subgroup meta-analysis in such a way that interactions between moderators that can account for variability in effect sizes derived from primary studies can be discovered. Dusseldorp et al found a number of effective combinations of BCTs for promoting physical activity and healthy eating: (1) "Provide information about behavior-health link" with "Prompt intention formation" (mean effect size =0.46), and (2) "Provide information on consequences" and "Use of follow-up prompts" (=0.44) [19]. However, little is known about synergistic effects in online interventions, which often include several BCTs. This research focuses on BCTs that have a cumulative effect in addition to their univariate positive effect. The second aim of this research is to identify synergistic effects of BCTs (next to their already identified univariate effects) in online interventions aimed at health behavior change.

A third factor that may influence effect sizes is the mode by which the intervention is delivered. Online interventions can differ substantially in their specific modes of delivery (MoD). Webb et al [3] noted that content can be delivered in a more or less interactive manner [18] and that online interventions that employ supplementary delivery modes (notably, text messaging, tailored feedback, access to advisor, telephone, or email) tend to be more effective in univariate analyses [3,10]. However, until now, research has not considered the effectiveness of different combinations of MoDs and how MoDs might interact with the use of particular BCTs and usability factors to determine the efficacy of an intervention. For example, an online peer forum (an MoD) is more likely to provide quick social support (a BCT), and not only at one weekly face-to-face session. Therefore, the third and fourth aims of this research are to identify the most effective combinations of MoDs and the most effective combinations of BCTs, MoDs, and usability factors, respectively. Providing insight into the effects of usability factors and synergistic effects with BCTs and MoDs will also provide a starting point for an evidence-based instrument that can be used to develop new interventions and evaluate the quality and the potential of existing interventions.

Our Research

This review aims to develop a taxonomy for coding the usability of online interventions and to identify what combinations of BCTs, MoDs, and usability factors influence the effectiveness of online interventions designed to promote health-related behavior. In order to identify these synergistic effects, meta-CART analysis was employed.

Methods

We considered data from the 85 studies that were included in the meta-analysis of Webb et al [3] for inclusion in this meta-analysis. The studies were published between 1990 and July 2008, in peer-reviewed journals and conference proceedings written in English. This review uses data on the effectiveness of the interventions and the use of BCTs and MoDs. Each intervention was coded by Webb et al for inclusion or exclusion of each of the 40 BCTs from the CALO-RE (Coventry, Aberdeen, and London—Refined) taxonomy of Michie et al [21]. Webb et al also coded 11 modes of delivery used by each of the interventions. Study characteristics were coded by a single

author and so information is not available on the reliability of so doing. Evidence suggests that behavior change techniques can be reliably coded [21], but additional research is needed to confirm the extent to which coders can reliably identify modes of delivery and usability.

Selection of Interventions

Webb et al included studies (1) in which the described intervention was delivered via the Internet, (2) where participants were randomly assigned to conditions, and (3) where health-related behavior was measured after the intervention. This review included interventions that used two or more BCTs, MoDs, or usability factors (if studies included only a single BCT, then it was not possible to study the combined effects of such factors). Second, only effective factors were included in the analyses, because initial meta-CART analyses found that including all factors resulted in a tree without any boosting (or strengthening) effects. A BCT or MoD was considered effective when univariate analyses showed that studies including the BCT or MoD had a higher absolute effect size than the studies not using the specific BCTs or MoD. Because the univariate effects of the usability factors still needed to be investigated, we based our selection only on the effects of the BCTs and MoDs.

A total of 52 studies met the criteria for inclusion in the review. The included studies reported on interventions targeting health-related behavior, such as physical activity (n=13), dietary intake (n=8), and alcohol consumption (n=6). Seven studies addressed multiple behaviors (eg, combined physical activity + dietary intake). The target population varied from children to adults and from the general population to patients at specific risk (eg, patients with diabetes). In order to evaluate the effect of the interventions, all of the studies compared an experimental condition that was exposed to the intervention with an active or passive control condition. The included studies tested the impact of 20 effective BCTs and 7 effective MoDs (see Multimedia Appendix 1 for an overview of all included BCTs and MoDs).

Taxonomy of Factors That Influence the Usability of an Intervention

Given that studies generally lack a description of the usability of the evaluated intervention, we developed a survey to obtain this information from the original authors. The survey was based on a taxonomy, which described indicators of the usability of online interventions. The existing literature points to various factors that influence how usable an intervention is likely to be, such as guidelines on functionality [22], accessibility [23],

usability [24], design [25], user experiences [26], as well as studies of persuasive technology [27] and Shneiderman's golden rules [28]. First, it was decided which factors were applicable for online interventions targeting health behaviors. Those factors were included in our first draft survey. Second, factors in the draft were grouped and summarized where possible. Third, titles and definitions were adapted to fit the human-computer interaction. Fourth, the factors and definitions were discussed with other interaction specialists. When necessary, factors were adapted or refined and new factors were added.

The result of these steps was a 27-item taxonomy, with 8 subscales: learnability (reflecting how easy it was to accomplish basic tasks the first time that a user encounters the design), efficiency (the speed at which tasks can be performed once users have learned the design), memorability (reflecting how easy it was to re-establish proficiency when a user returns to the design after a period of not using it), errors (likelihood and severity of errors potentially made by users, and how easy it was to recover from such errors), satisfaction (pleasantness of using the intervention), personalization (reflecting the extent to which it was possible to adapt to the intervention to the individual user's characteristics, preferences, values and self-image), situatedness (the ability of the system to predict and adapt to the user's dynamic behavior in specifics contexts), and social interaction (reflecting whether the intervention encouraged social interactions). These features and the attributes that compose them are shown in Table 1.

We contacted the 85 authors of the original papers and asked them to rate their intervention using a questionnaire version of our taxonomy and to send us their intervention (if available) so that we could also code features influencing usability. Twelve authors (14%) coded their intervention and six (7%) provided access to their intervention or pictures from the intervention. Calculations and analyses are based on these 12 studies. We calculated scores for each study on each of the 8 scales of the taxonomy by calculating the mean of the items in each scale. Descriptive statistics are used to calculate the scale reliabilities, means, and standard deviations. Scale reliability varied from Cronbach alpha=.69 (efficiency) to .98 (situatedness). The learnability scale had an alpha of only .52 (even after omitting an item) and was therefore not included in the analyses. Scale scores were then dichotomized based on the median (ie, low or high on the scale). Finally, the effect sizes for the low and high group were calculated for each usability factor by using subgroup analyses in Comprehensive Meta-Analysis (CMA) software, version 2.2 [29].

Table 1. Usability taxonomy (the intervention referred to in the taxonomy is the intervention itself as well as its embodiment, the user interface).

Usability attributes	Measures (example)
Learnability	
1. Consistency	The intervention is consistent in the use of interface aspects such as layout, buttons, and language (eg, the OK button is always on the left and the Cancel button is always on the right side, consistent use of terminology and look-and-feel)
2. Conventions	The intervention follows platform conventions (eg, in Windows, the cross at the upper right corner of the screen is always used to close the window)
3. Intuitiveness	The intervention characteristics intuitively imply its functionality and use (eg, a button with an arrow pointing to the right, implying "go to the next page")
4. Visibility of system status	The intervention provides feedback about its (future) state, action, and result (eg, when loading, the system provides a load bar showing how much time has passed and how much time remains)
Efficiency	
5. Flexibility	The intervention caters to a variety of users, both inexperienced and experienced (eg, the system provides both viewable icons, such as a floppy disk, and short-cuts, such as Ctrl-S)
6. Structure	Using the intervention, users understand the structure of the intervention and know where they are (eg, the intervention provides breadcrumb navigation, ie, showing previous interaction steps and steps to come)
7. Defaults	The intervention makes use of default settings (eg, fields containing defaults come up selected and the user can replace the default contents with new information—the defaults are user-specific)
Memorability	
8. Tailoring to user group	The intervention speaks the user groups' language, with words, phrases and concepts familiar to the user group, rather than intervention-oriented terms (eg, intervention contains words and phrases fit for children, intervention uses read-aloud function for low literates)
9. Recognition rather than recall	The intervention minimizes the user's effort by making options visible or easily retrievable whenever appropriate (eg, the intervention shows context specific relevant available functionalities instead of referring to a manual where all options are listed)
Errors	
10. Error recovery	The intervention supports undo and redo (eg, the intervention offers a "Go" and a "Back" button)
11. Error prevention by the system	The intervention prevents problems from occurring and notifies the user if a problem can potentially occur (eg, the intervention indicates which fields are mandatory [*] and applies form validation, such as the right format for postal code)
12. Error recognition and resolution by the system	The intervention provides error messages expressed in plain language (no codes), which precisely indicate the problem and constructively suggest a solution (eg, when entering a faulty password, the intervention indicates: "Your password is incorrect, please ensure your CAPS LOCK key is off")
13. Help by the system	The intervention provides help information that is easy to search and focuses on the user's task (eg, the intervention has a help function, in the form of a question mark icon, for every text field that needs to be entered)
Satisfaction	
14. Minimalistic	The intervention does not contain elements that are irrelevant or rarely needed (eg, the intervention interface is not cluttered, does not use distracting irrelevant interface elements, and does not require extensive scrolling)
15. Aesthetic appearance	The intervention is esthetically attractive (eg, includes pictures, colors)
16. Fun	The intervention is fun to use (eg, the intervention offers a challenge to users and arouses their curiosity)
17. Modality integrity	The intervention offers information in a suitable modality (eg, information is presented in text when presenting details, information is presented in image when providing an overview)
18. User control	The user is in control of the intervention (eg, the user initiates actions, the system justifies responses to actions of the user)
Personalization	
19. Adaptability by the user	The user can adapt the intervention to fit to their preferences and skill level (eg, the intervention offers the user to decrease and increase interface's font size, the intervention offers the possibility to take a tour through the intervention or change the background color)
20. Adaptiveness to the user	The intervention is aware of the user's characteristics and adapts the interface to these characteristics (eg, the intervention is aware the user is farsighted and increases the font size)
21. Adaptiveness to the context	The intervention is aware of the user's context and adapts the interface to this context (eg, when the user is using the intervention in a public space, sound is turned off)

XSL•FO
RenderX

Usability attributes	Measures (example)
22. Adaptiveness to the task	The intervention is aware of the task that the user aims to perform and adapts the interface to this task (eg, the intervention detects the user is making a presentation and the intervention provides the tools generally used, such as inserting text boxes and images)
Situatedness	
23. Prediction of behavior	The intervention accurately predicts the user's behavior (eg, the intervention detects the user is in the office at 4 p.m. and predicts they will probably go home within the hour)
24. Adaptation to behavior	The intervention accurately adapts the interface to the predicted behavior (eg, the intervention predicts they will go home within the hour and automatically offers a weather and traffic forecast)
Social interaction (these three questions apply to the automated system, and not to contact with a human person such as a therapist or health care professional)	
25. Embodiment	The online intervention appears humanlike (eg, the intervention has a face, eyes, and/or body that are used for non-verbal communication. This can be in the form of a robot, avatar, or character)
26. Communication	The intervention applies humanlike communication skills (eg, the intervention expresses social thoughts, feelings, and behaviors)
27. Following conventions	The intervention follows and applies social conventions (eg, the intervention applies turn taking, cooperation, and/or reciprocity)

Outcome Measures

Following Webb et al [3], when studies reported the impact of an intervention on multiple outcomes, effect sizes were averaged before inclusion in the main dataset. Where studies included multiple points of measurement, the longest follow-up point was included. Effect sizes were computed as the standardized mean difference between intervention and comparison conditions in study outcomes using Hedges' correction for small sample size [30]. In our study, the distribution of the effect sizes was dichotomized using the overall effect size as the split point (following the strategy proposed by [19]). The effect sizes of the 52 included interventions ranged from -0.47 to 2.25, with an overall pooled effect size (weighted for sample size) of =0.25 (95% CI 0.18-0.31). The median value was =0.20, which is considered a small effect size, according to Cohen's criteria [31]. We used this latter value as a criterion for success for this type of intervention. Interventions with an effect size higher or equal to =0.20 were classified as successful (n=26), and those with effect sizes below 0.20 as less successful (n=25).

Statistical Analysis

Prior to the main analyses, effect sizes were inspected for outliers. The intervention described by Hurling et al [32] appeared to be an outlier (=2.25). Analyses performed with and without this outlier showed that it significantly influenced the weighted average effects and so it was omitted from subsequent analyses.

Univariate Analyses

To analyze the effect of usability factors on the efficacy of interventions, standard subgroup or moderator analyses, were conducted. The mixed effects model consisted of a random effects model within subgroups and a fixed effect model *across* subgroups, which is an approach recommended by Borenstein et al [33]. The significance of the Q model statistic indicated whether the heterogeneity could be explained by the between groups variable and thus if the factor was a significant moderator [33]. The mixed effects analysis was performed in CMA [29].

Multivariate Analyses With Meta-CART

Meta-CART consists of two phases. In the first phase, a CART analysis is applied and in the second phase, subgroup meta-analysis is applied to the results of the first phase. CART is a machine learning technique that builds classification trees for categorical outcome variables and regression trees for continuous outcome variables. In the context of our review, the CART algorithm partitions interventions into homogeneous subsets, resulting in a binary tree in which the end nodes contain the most homogeneous groups with respect to within-group effect size. The partitioning is based on intervention characteristics (eg, a BCT). More information on the background of CART analysis is provided by Dusseldorp et al [19].

A CART analysis proceeds in three steps (see Figure 1). In the first step, a full classification tree is grown [20] for the dichotomized outcome variable with different minimum numbers of interventions in an end node. This minimum can be fixed at 5 (which is often used in standard CART) or can be varied. In general, the recommended strategy is to choose this value as low as possible (to be able to grow a large tree and then prune it back) [20]. In this study, we varied this number between 2 (the absolute minimum) and 6 (a relatively high value). In addition, a minimal decrease in heterogeneity (impurity) of 0.001 was set as a stopping rule. In this analysis, this resulted in 6 full classification trees that differed in the minimum number of interventions in the end nodes. In the second step, the full classification trees were pruned, using the standard procedure of CART (with tenfold cross-validation and the one-standard-error rule [19]). The pruning procedure results in a best size of the tree, expressed in the number of end nodes. To increase the stability of the results, the pruning procedure was repeated 1000 times [19]. This resulted in 1000 estimates of the best tree size, from which the modal tree size was chosen. If the modal tree size was 1 (meaning a tree of one end node, ie, the total group is not split) or 2 (meaning a tree with two end notes, so one split, indicating no interaction), then the analyses were not continued. For example, in Figure 1, the pruning procedure shows that there is no interaction effect when the

minimum number in an end note is set at 2. In the third step, the pruned trees were inspected for "end cut preference" splitting: this occurs if the first split of the tree ends in a node that has the minimum number of interventions (as defined earlier, ranging from 2-6). If this occurs in the first branch, no splits can be made after this one. If end-cut preference was present, the tree was dismissed and a tree with a larger minimum number of interventions was preferred [34]. In this research, end-cut preference occurred when the minimum number was set at 3, 4, and 5. These three steps are shown in Figure 1. The final tree represents the synergistic effect of the moderators (in our case BCTs and or modes) on outcomes. The end nodes of the tree form the subgroups. From the final tree, a new variable was created, with its categories referring to the end nodes of the tree. The CART analyses were performed in the R software environment, version 2.15 [35] using the package rpart, which is developed for classification and regression trees [36].

In the second phase of the meta-CART procedure, a standard subgroup meta-analysis is performed to investigate whether the new grouping variable resulting from the first phase accounts for heterogeneity in the study effect sizes. The same procedures were used as for the analyses of the factors that influence usability. An advantage of the subgroup analysis was that a weighted mean effect size ( ) was obtained for each subgroup (ie, end node of the tree).

In total, three meta-CART analyses were performed, varying in the BCTs and MoDs that were included as moderators (see Table 2). The first analysis included only the 20 BCTs. The second analysis included only the 7 MoDs. The third analysis included the 20 BCTs and 7 MoDs together. Not enough studies were available to perform a meta-CART analysis with the usability factors as moderators.

Table 2. Meta-CART Analyses conducted in this research (see Multimedia Appendix 1 for the names of the BCTs and MoD).

Research question	Included moderators
	BCT/ MoD[a]
Effective combinations of BCTs	Analysis 1: 46 studies
	20 BCTs: 4, 5, 6, 7, 8, 10, 13, 14, 19, 20, 21, 22, 23, 24, 25, 27, 28, 34, 35, 39
Effective combinations of MoD	Analysis 2: 31 studies
	7 MoDs: b, d, e, f, g, h, i
Effective combinations of BCTs and MoDs	Analysis 3: 43 studies
	20 BCTs (see above); 7 MoDs (see above)

[a]BCTs/ MoDs that are univariately associated with an effect size that is higher than the effect size of the studies that do not include the specific BCT/MoD, minus outlier Hurling et al [32].

Results

First, the scale characteristics and effect sizes from the factors that influence usability are described. Then, the results from the three meta-CART analyses are described: the first one using BCTs only, the second one using MoDs only, and the third one using both.

Usability and Effect Size

The first analysis examined whether the usability factors influenced intervention effectiveness. Interventions that were deemed to be more efficient (ie, scored higher on the efficiency subscale) proved to be more effective (=0.43, 95% CI 0.18-0.67) than interventions that are less efficient (=0.02, 95%CI -0.29 to 0.10; between groups Q-value=4.25, *P*=.04). No relations between effect size and other usability factors were found (see Table 3).

Figure 1. Stepwise meta-CART analyses.
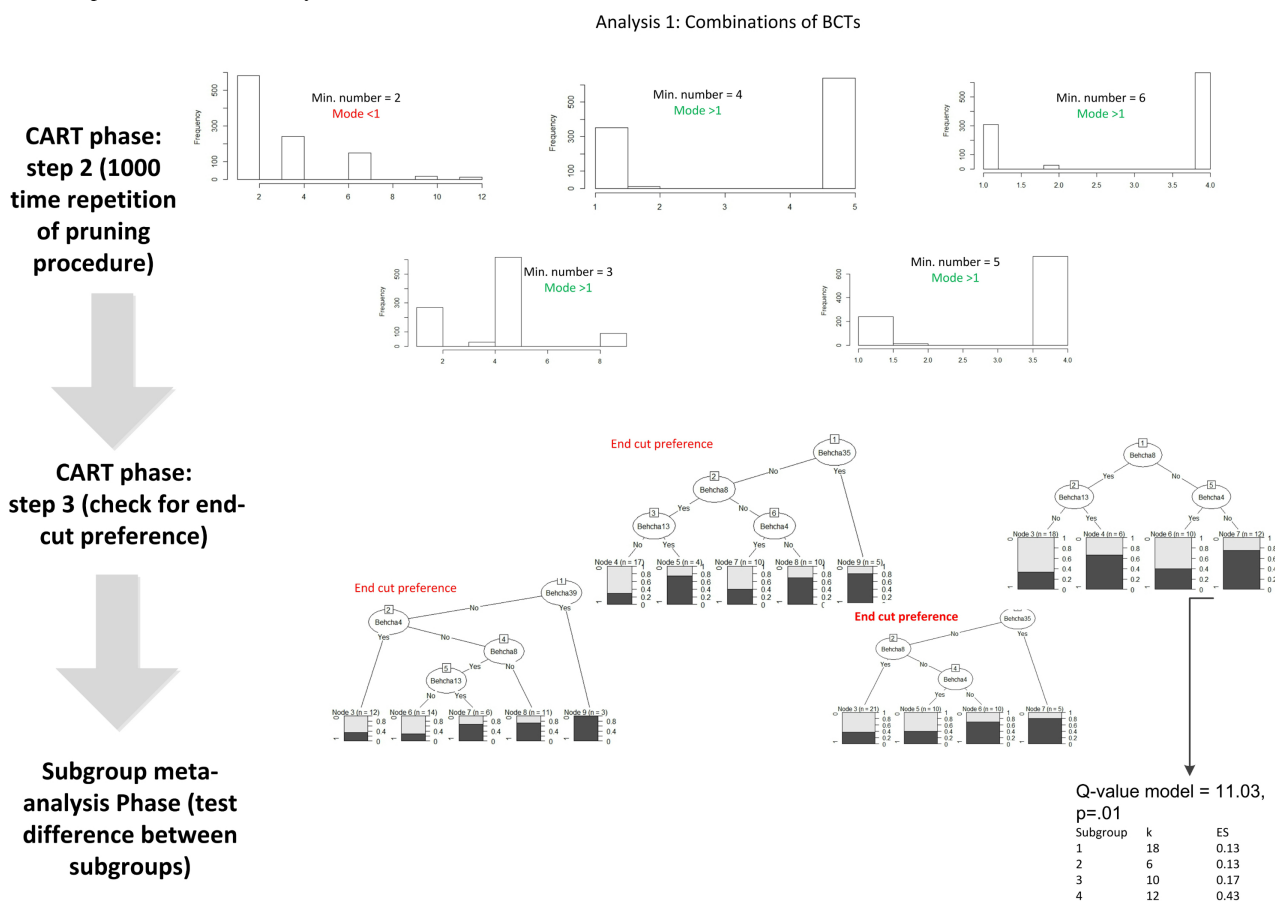
Table 3. Usability: factor reliabilities, factor median, and the impact of these factors on effect size (low versus high, Q-model and *P* value for difference) among 12 interventions.

Scale	α	Median of scale	(95% CI)		Between groups Q (*P* value)
			Low (≤median of scale)	High (>median of scale)	
Learnability[a]	.52[b]	-	-	-	-
Efficiency	.69	3.33	0.02 (-0.29 to 0.10)	0.43 (0.18-0.67)	4.25 (.04)
Memorability	.91	5.00	0.32 (-0.23 to 0.87)	0.22 (-0.09 to 0.53)	0.09 (.76)
Errors[c]	.70	3.33	0.47 (0.31-0.63)	0.15 (-0.34 to 0.64)	1.45 (.23)
Satisfaction[d]	.73	4.00	0.37 (0.01-0.73)	0.15 (-0.22 to 0.53)	0.67 (.41)
Personalization[e]	.70	2.33	0.15 (-0.19 to 0.49)	0.38 (0.05-0.68)	0.82 (.37)
Situatedness	.98	2.00	0.21 (-0.08 to 0.50)	0.52 (0.23-0.81)	2.14 (.14)
Social Interaction	.94	3.33	0.10 (-0.40 to 0.69)	0.34 (0.04-0.65)	0.68 (.41)

[a]Item 4 removed from scale.

[b]Scale not analyzed because of low reliability.

[c]Item 10 removed from scale.

[d]Item 16 removed from scale.

[e]Item 21 removed from scale.

Effective Combinations of Behavior Change Techniques

Meta-CART analyses were conducted to study effective combinations of BCTs. The best fitting tree was found when a BCT was used by at least 6 interventions, that is, the minimum amount of studies in an end node was set at 6. This resulted in a tree with four end nodes (see Figures 1 and 2). Subgroup analysis showed that the subgroups were significantly different from each other (between groups Q-value=11.03, *P*=.01). One synergistic effect was found: interventions that included both barrier identification/ problem solving and provided rewards

for behavior change had an average effect size of 0.23 (95% CI 0.08-0.38). Interventions that only provided normative information about the behavior of others (=0.16, 95% CI 0.10-0.23) or included barrier identification/ problem solving

(=0.13, 95% CI 0.04-0.23) were least effective. Interventions that used other combinations of techniques than barrier identification/ problem solving and provided rewards for behavior change were most effective (=0.43, 95% CI 0.27-0.59).

Figure 2. Results from meta-CART and subgroup analyses: classification tree and effect sizes across studies that used at least two univariately effective BCTs (n=47) (note a: percentage of interventions in this node that were more successful, ie, an effect size higher than 0.20).
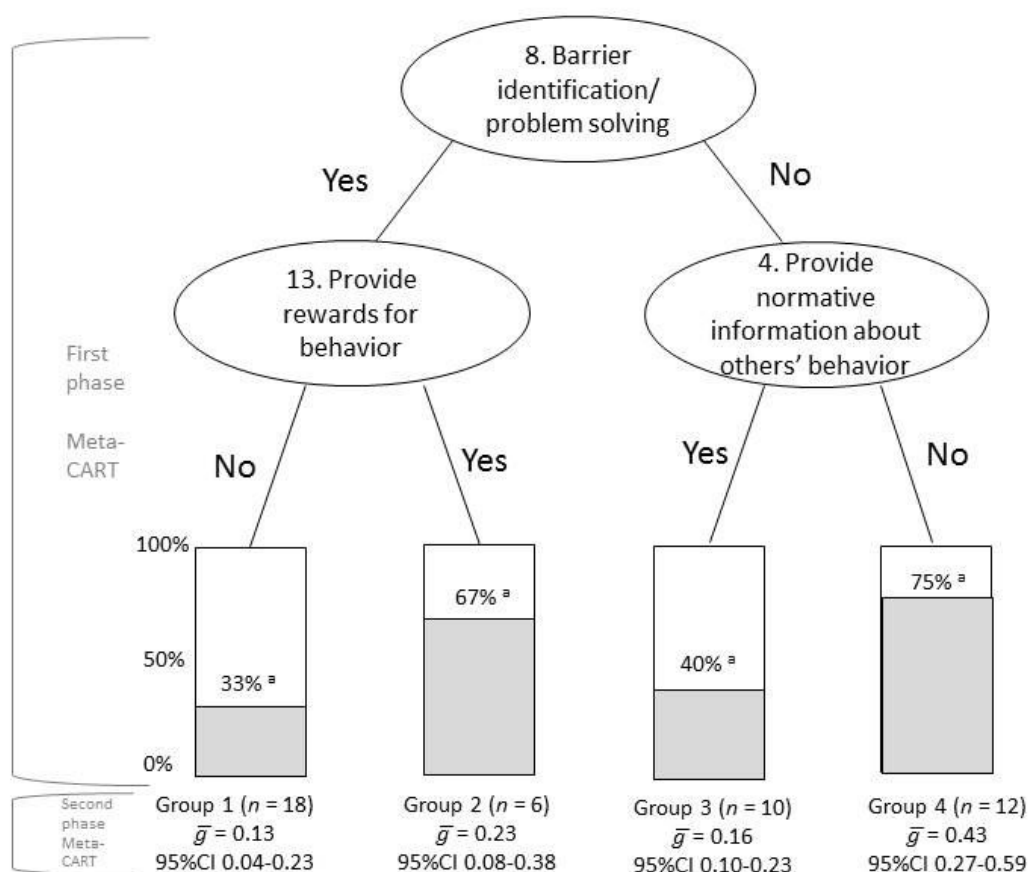
Effective Combinations of Modes of Delivery

The meta-CART analysis including the seven MoDs showed that the best fitting tree had only one node—the root node. In other words, there was no combination of MoDs that was able to explain heterogeneity in the effectiveness of Internet-based interventions for health behavior change.

Effective Combinations of Behavior Change Techniques and Modes of Delivery

When all BCTs and MoDs were combined in the third analysis, no synergistic effects were found.

Additional Analyses on Group 4

Figure 2 shows that interventions that did not prompt barrier identification/ problem solving or provide normative information about others' behavior (Group 4) were the most effective. We therefore conducted additional analyses to understand the

positive effects among this group of studies. Based on previous reviews, meta-analyses and the original study by Webb, we set hypothesis related to study quality, intervention intensity, and effective univariate BCTs. Our first hypothesis was that the increased effectiveness was due to the interventions in this group being more intensive. Our second hypothesis was that the interventions with a larger effect were evaluated in studies of lower methodological quality than the interventions with a smaller effect (eg, used passive control conditions or self-report measures of outcome). In order to test these hypotheses, 2 authors coded (1) the number of contacts and length of contact in the intervention, (2) activity in the control condition (no treatment vs alternative intervention), and (3) measurement of outcome (self-report or objective). They discussed the coding results until agreement was reached.

Next, we compared the BCTs, intervention, and characteristics from the studies in Group 4 (n=12) to the other studies in this

review (n=34), using chi-square tests and *t* tests. Interventions in Group 4 had significantly more contact moments (mean=180) than the other interventions (mean 40; *P*=.03). Furthermore, the BCTs "model/demonstrate the behavior" (*P*=.09) and "prompt self-talk" (*P*=.01) were more often used in Group 4 studies compared to the other studies. These BCTs were not included in the meta-CART because they were used by fewer than 6 studies. The interventions in Group 4 also used an active control intervention more often (compared to no intervention, *P*=.03) and had a longer follow-up time between intervention and measurement (*P* <.05) than the other studies. No differences were found with regard to use of self-reporting or objective measurement of outcomes. Thus, the higher effects in this group cannot be explained by a weaker study design in these studies.

Discussion

Principal Findings

This meta-analysis re-analyzed data from a systematic review of online interventions aimed at health-related behavior change [3] in an effort to identify synergistic effects of behavior change techniques (BCTs), modes of delivery (MoDs), and usability factors. First, the univariate analyses of the usability factors indicated that one usability factor influenced effect sizes: the efficiency of the intervention. Meta-CART was then used to identify subgroups of interventions that were associated with particular levels of effectiveness. One synergistic effect was found: barrier identification/ problem solving in combination with providing rewards for behavior change (=0.23). No synergistic effects were found among MoDs or among MoDs with BCTs. Below, we discuss each of these findings in more detail.

Usability Factors

The usability questionnaire that we developed resulted in 7 reliable subscales (one subscale was not reliable). Efficiency was positively related to intervention effect. This means that interventions that are flexible (can cater both experienced and inexperienced users), provide structure (that is understood by the user and the user knows where they are in that structure), and make use of default settings are more likely to be effective. The score for efficiency ranged from 2.33 (for study [37]) to 5 (for study [38,39]). Participants may be more likely to use an intervention if it is efficient in its use (eg, quick delivery of information or support). Furthermore, these characteristics allow the user to use the system easily and quickly, therefore posing low cognitive load on the user. A high cognitive load is less likely to result in a learning experience than a low cognitive load [40]. Based on these assumptions, increasing the content of an online intervention may come at the expense of efficiency. This makes it necessary to take into account usability and effectiveness when making a strategic choice about what to include in interventions.

Nevertheless, we should be careful in interpreting these results because usability information was provided by the authors of only 12 of the primary studies and their answers were not cross-validated as we did not have access to these interventions. However, based on the reliability of the subscales (the majority) and the finding that efficiency predicted variability in the

effectiveness of the primary interventions, we believe that our taxonomy is suitable for use in these kinds of settings. Of course, more research in larger studies is needed to establish its reliability and validity, especially if coded by different authors. A special concern in these studies should be the improvement of the learnability scale, as its reliability was unsatisfactory. Better-powered studies should explore the best combinations of items to make the scales. In addition, item 4 (visibility of system status) was removed because it reduced the reliability of the scale even further. The remaining three items have a mean value of 4.4-4.6 (SD 0.70-0.97). As the minimum is 1 and the maximum is 5, this may indicate that these questions cannot differentiate sufficiently. Perhaps the questions should be stated in a stronger and more precise fashion. For example, instead of asking if the intervention follows platform conventions (item 2; scale of totally disagree to totally agree), it may be possible to ask which or how many aspect platform conventions are followed (eg, cross in upper right corner, disk to save changes).

Synergistic Effects

Webb et al [3] reported that the univariate effects of barrier identification/ problem solving and providing rewards for behavior change were =0.20 and =0.18, respectively. In this analysis, these two BCTs had a combined effect of =0.23, which is higher than using barrier identification/ problem solving only. Both techniques are likely to play a role in the maintenance of action and prevention of relapse. The interactive effect may be due to their different approach. While barrier identification/ problem solving takes a somewhat negative approach (ie, thinking about what can go wrong), providing rewards for behavior change suggests a positive consequence in the future. Also, barrier identification/ problem solving is aimed directly at problem solving (instrumental function of skill learning), while providing rewards for behavior change has a more affective function in increasing motivation. The combination may thus provide for different needs. Based on our findings, we suggest using these techniques together.

However, the most effective interventions were those that did not provide normative information about the behavior of others or prompt barrier identification/ problem solving (=0.43). Webb et al [3] showed that interventions including more BCTs were more effective, but this group of studies did not use more BCTs than the less effective studies. However, the effective studies had significantly more contact moments than the other studies. In addition, these studies more often used "model/demonstrate the behavior" and "prompt self-talk" than the other studies. Furthermore, it is likely that these studies (and also the other studies) included BCTs that were not included in the taxonomy. Thus, they are not coded and their effects not analyzed by Webb, nor by us in this study. Also, implementation of the BCTs was not taken into account, a factor that may influence effectiveness [41] by how, when, and for how long a BCT is used in an intervention. For example, the BCT "provide rewards for behavior" does not make explicit what the reward is, who decides it, and how often it is given. These factors can greatly influence its effectiveness.

In line with previous studies, our findings suggest that online interventions designed to promote health behavior should not

provide normative information on the behavior of others. Other studies showed that providing information on what people usually do (ie, descriptive norms), independent of presenting minority or majority information [42], is less likely to result in action [39,43,44]. When presenting normative information, it might be better to use injunctive norms (what people typically approve or disapprove) [45]. In future reviews it will be important to differentiate between injunctive and descriptive norms and then analyze differential effects.

It is important to understand the context in which a BCT operates [42-44]. MoD is part of this context. Michie et al [18] showed that differences in the effectiveness of the interventions could be explained by, among others, theoretical basis, use of theory and MoD. In the original review [3], the univariate influence of additional MoD was quite strong (eg, Internet-based interventions that also included telephone contact were more effective than those without). In our research, however, no effective combinations were found among MoD. This may be due to the fact that each MoD by itself is already an additional mode: every mode is in addition to an online intervention. Based on the results of Webb et al's review and our meta-analysis, we can recommend that online interventions include one or more additional MoDs, but we cannot advise on specific additional modes to increase the effectiveness. In addition, when BCTs and MoDs were combined, only combinations of BCTs were found. These findings indicate that the effectiveness of BCTs is not dependent on (additional) modes of delivery.

Limitations and Future Directions

This is one of the first studies to investigate interactive effects of BCTs and MoDs in online interventions. However, the lack of primary studies that used a large number of BCTs and MoDs limited our options for analysis. First, not all BCTs from the original taxonomy of Abraham and Michie [9] were present in the original analyses by Webb et al [3]. Furthermore, only a selection of the BCTs from Webb et al was included in the meta-CART analysis. In addition, other techniques that are not part of the taxonomy (such as cognitive restructuring) may be effective as well. A similar limitation occurred for modes of delivery. As such, our findings are limited to the relatively small number of effective BCTs and MODs that were present in at least 6 interventions (the number needed in each end node to obtain a stable tree) in this dataset. Stress management, general communication skills training, model/demonstrate the behavior, and facilitate social comparison were highly effective univariately but were used by five studies or fewer. As such, these BCTs were not included in the final tree. Thus, the lack of effects found for these BCTs in this meta-analysis does not influence the conclusions about these BCTs in [3]. Furthermore, all studies in the meta-analysis were published at least 7 years ago. As such, no interventions using smartphones or tablets have been included. It would be interesting to study if the same effects are found by more recent interventions, and if the same relation with usability exists.

In general, few synergistic effects were found, suggesting that the BCTs that were identified by Webb et al [3] have a clear single effect that cannot [46] be strengthened by combining them with other BCTs. For example, previous studies have shown that self-monitoring of behavioral outcomes is moderately-to-strongly related to behavior change [3,18,47-50]. In the current meta-analysis, we could not identify other BCTs that improve the effects of self-monitoring. The lack of this combined effect may be due to the same reason why we did not find any combination among some of the most effective BCTs: the low number of studies that used a certain BCT. Another problem may be the dichotomization of the effect size (using a classification tree). Ideally, the original, continuous outcome would be used (in a regression tree); however, the studies could not be weighted. In the CART procedure, the accuracy of the study effect sizes could not be taken into account. The advantage of the use of a continuous outcome (compared to a dichotomized one) is that the variance in the effect sizes is maintained. Therefore, a regression tree is less likely to be biased than a classification tree. The next step for CART in meta-analysis is to develop the possibility of taking study weight into account, thus decreasing the risk of bias in trees.

This meta-analysis aimed to provide information on which combination of techniques may enhance the effectiveness of Internet-based interventions designed to promote health behavior. It does not, however, explain how the techniques should be used in practice, that is, when, where, and in which shape they will be effective. Bartholomew et al provide such recommendations for a limited set of techniques [51], but more information is needed. Also, negative effects were not explored further in the meta-CART analyses because interactions were found only when factors that have positive effects were included. Another limitation is that this meta-analysis shows the effects on behavior and/or outcomes, but not on the potential mediators (eg, factors such as attitudes, self-efficacy, and skills) that might explain behavior change [46,52]. Given that behavior change is assumed to be the result of changes via behavioral or environmental determinants [21,53], and BCTs are expected to have an impact on such determinants, it is important to also understand the change mechanisms [54]. More evidence is also needed with regard to implementation factors such as dose and fidelity [13,55,56]. A higher number of intended contact moments was associated with a larger effects size in the Group 4 studies, but we did not take into account the intended or delivered dose of BCTs. Adding a measure of actual dose allows us to study a dose-effect relation and differentiate between efficacy and effectiveness [41] and may thus increase our understanding of when and how to choose and use BCTs.

Further research into usability factors may also increase our insight in change mechanisms. The lists of indicators of usability and mode of delivery that were used in this study are not exhaustive. Future studies should take into account recent developments in this field, such as the use of smartphones and their connection with the intervention under study. Finally, a broad range of outcomes was used in this meta-analysis (physical activity, healthy eating, alcohol consumption and more) and the relatively small number of studies did not allow us to conduct separate analyses for each behavior. In future studies, it would be useful to find out if combinations of BCTs have different effects for specific behaviors.

XSL•FO
RenderX

Conclusion

This research developed a new coding frame for identifying the factors that influence the usability of online interventions. One factor—efficiency—influenced effect sizes, with more efficient studies tending to report larger effects than less efficient studies.

We then used meta-CART analysis to investigate the effects of combinations of BCTs and/or modes of delivery. We were able to identify the synergistic effect of two BCTs: (1) prompting barrier identification/ problem solving in combination with (2) providing rewards for behavior change. However, no other interactive effects of BCTs and modes of delivery were found.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Univariate effect sizes of behavior change techniques and modes of delivery (n=85 studies).

[PDF File (Adobe PDF File), 51KB - jmir_v5i2e61_app1.pdf]

References

1. Fox S. Pew Research Center's Internet & American Life Project. 2011. The social life of health information URL: http://www.pewinternet.org/~/media/Files/Reports/2011/PIP_Social_Life_of_Health_Info.pdf [accessed 2015-01-09] [WebCite Cache ID 6VRoNyr0R]

2. 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]

3. 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]

4. 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]

5. Shahab L, McEwen A. Online support for smoking cessation: a systematic review of the literature. Addiction 2009 Nov;104(11):1792-1804. [doi: 10.1111/j.1360-0443.2009.02710.x] [Medline: 19832783]

6. Verweij LM, Coffeng J, van MW, Proper KI. Meta-analyses of workplace physical activity and dietary behaviour interventions on weight outcomes. Obes Rev 2011 Jun;12(6):406-429. [doi: 10.1111/j.1467-789X.2010.00765.x] [Medline: 20546142]

7. Abraham C, Sheeran P, Johnston M. From health beliefs to self-regulation: Theoretical advances in the psychology of action control. Psychology & Health 1998 Jul;13(4):569-591. [doi: 10.1080/08870449808407420]

8. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]

9. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]

10. 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]

11. Peters GY, Ruiter RAC, Kok G. Threatening communication: a critical re-analysis and a revised meta-analytic test of fear appeal theory. Health Psychol Rev 2013 May;7(Suppl 1):S8-S31 [FREE Full text] [doi: 10.1080/17437199.2012.703527] [Medline: 23772231]

12. Sheeran P, Harris PR, Epton T. Does heightening risk appraisals change people's intentions and behavior? A meta-analysis of experimental studies. Psychol Bull 2014 Mar;140(2):511-543. [doi: 10.1037/a0033065] [Medline: 23731175]

13. Breiman L. Classification and regression trees. New York: Chapman & Hall; 1984.

14. Dusseldorp E, van GL, van BS, Verheijden MW, van EP. Combinations of techniques that effectively change health behavior: evidence from Meta-CART analysis. Health Psychol 2014 Dec;33(12):1530-1540. [doi: 10.1037/hea0000018] [Medline: 24274802]

15. Brouwer W, Kroeze W, Crutzen R, de NJ, 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]

16. Crutzen R, de Nooijer J, Brouwer W, Oenema A, Brug J, de Vries NK. Strategies to facilitate exposure to internet-delivered health behavior change interventions aimed at adolescents or young adults: a systematic review. Health Educ Behav 2011 Feb;38(1):49-62. [doi: 10.1177/1090198110372878] [Medline: 21189422]

XSL·FO
RenderX

17.   van Genugten L, van Empelen P, Boon B, Borsboom G, Visscher T, Oenema A. Results from an online computer-tailored weight management intervention for overweight adults: randomized controlled trial. J Med Internet Res 2012;14(2):e44 [FREE Full text] [doi: 10.2196/jmir.1901] [Medline: 22417813]

18.   Hurling R, Fairley BW, Dias MB. Internet-based exercise intervention systems: Are more interactive designs better? Psychology & Health 2006 Dec;21(6):757-772. [doi: 10.1080/14768320600603257]

19.   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]

20.   Nielsen J. Designing Web Usability : The Practice of Simplicity. San Francisco: New Riders Press; 2000.

21.   Michie S, Ashford S, Sniehotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. Psychol Health 2011 Nov;26(11):1479-1498. [doi: 10.1080/08870446.2010.540664] [Medline: 21678185]

22.   Davis FD. User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. International Journal of Man-Machine Studies 1993 Mar;38(3):475-487. [doi: 10.1006/imms.1993.1022]

23.   Allan K, Ford K, Patch K, Spellman J. User agent accessibility guidelines. URL: http://www.w3.org/TR/UAAG20/ [accessed 2015-01-07] [WebCite Cache ID 6VP0ao66Q]

24.   Nielsen J, Mack RP. Usability inspection methods. New York: Wiley; 1994.

25.   Rogers Y, Sharp Y, Preece J. Interaction Design: Beyond Human-Computer Interaction. New Jersey: Wiley; 2011.

26.   Park J, Han SH, Kim HK, Oh S, Moon H. Modeling user experience: A case study on a mobile device. International Journal of Industrial Ergonomics 2013 Mar;43(2):187-196. [doi: 10.1016/j.ergon.2013.01.005]

27.   Fogg B. Persuasive Technology: Using Computers to Change What We Think and Do. Burlington: Morgan Kaufmann; 2002.

28.   Shneiderman S, Plaisant C, Cohen M, Jacobs S. Designing the user interface: strategies for effective human-computer interaction. Boston: Addison-Wesley; 2010.

29.   Rothstein H, Sutton A, Borenstein M. Software for publication bias. In: Publication bias in meta-analysis: prevention, assessment and adjustments. Chichester, England: Wiley; 2005.

30.   Hedges L. Statistical Aspects of Effect Size Estimation. Washington: ERIC; 1981 Presented at: Annual Meeting of the American Educational Research Association; April 13-17, 1981; Los Angeles.

31.   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]

32.   Borenstein M, Hedges L, Rothstein H. Introduction to meta-analysis. Oxford: Wiley; 2009.

33.   Bradford J, Broadley C. The effect of instance-space partition on significance. Machine Learning 2001;42(3):269-286. [doi: 10.1023/A:1007613918580]

34.   R Development Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2012. URL: http://www.gbif.org/resource/81287 [accessed 2016-03-10] [WebCite Cache ID 6fuLmaGrK]

35.   Therneau T, Atkinson B, Ripley B. R package version. 2012. Rpart: recursive partitioning URL: https://cran.r-project.org/web/packages/rpart/index.html [accessed 2016-03-10] [WebCite Cache ID 6fuLvrwbN]

36.   Schrader C, Bastiaens T. The influence of virtual presence: Effects on experienced cognitive load and learning outcomes in educational computer games. Computers in Human Behavior 2012 Mar;28(2):648-658. [doi: 10.1016/j.chb.2011.11.011]

37.   Stok FM, de Ridder DTD, de Vet E, de Wit JBF. Minority talks: the influence of descriptive social norms on fruit intake. Psychol Health 2012;27(8):956-970. [doi: 10.1080/08870446.2011.635303] [Medline: 22149532]

38.   Burger JM, Shelton M. Changing everyday health behaviors through descriptive norm manipulations. Social Influence 2011 Apr;6(2):69-77. [doi: 10.1080/15534510.2010.542305]

39.   Jacobson RP, Mortensen CR, Cialdini RB. Bodies obliged and unbound: differentiated response tendencies for injunctive and descriptive social norms. J Pers Soc Psychol 2011 Mar;100(3):433-448. [doi: 10.1037/a0021470] [Medline: 21171790]

40.   Reid AE, Aiken LS. Correcting injunctive norm misperceptions motivates behavior change: a randomized controlled sun protection intervention. Health Psychol 2013 May;32(5):551-560. [doi: 10.1037/a0028140] [Medline: 23646838]

41.   Cialdini RB. Crafting normative messages to protect the environment. Current Directions in Psychol Sci 2003 Aug;12(4):105-109. [doi: 10.1111/1467-8721.01242]

42.   Rothman AJ, Baldwin AS, Hertel AW. Self-regulation behavior change: Disentangling behavioral initiationbehavioral maintenance. In: Baumeister RF, Vohs KD, editors. Handbook of Self-Regulation, Second Edition: Research, Theory, and Applications. New York: The Guilford Press; 2004.

43.   Kok G, Schaalma H, Ruiter RAC, van Empelen P, Brug J. Intervention mapping: protocol for applying health psychology theory to prevention programmes. J Health Psychol 2004 Jan;9(1):85-98. [doi: 10.1177/1359105304038379] [Medline: 14683571]

44.   Bartholomew L, Parcel GS, Kok G, Gottlieb NH, Fernandez ME. Planning Health Promotion Programs: An Intervention Mapping Approach. 3rd edition. San Francisco: Jossey-Bass; 2011.

XSL•FO

RenderX

45.   Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]

46.   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]

47.   Dombrowski SU, Sniehotta FF, Avenell A, Johnston M, MacLennan G, Araújo-Soares V. Identifying active ingredients in complex behavioural interventions for obese adults with obesity-related co-morbidities or additional risk factors for co-morbidities: a systematic review. Health Psychology Review 2012 Mar;6(1):7-32. [doi: 10.1080/17437199.2010.513298]

48.   Greaves CJ, Sheppard KE, Abraham C, Hardeman W, Roden M, Evans PH, et al. Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. BMC Public Health 2011;11:119 [FREE Full text] [doi: 10.1186/1471-2458-11-119] [Medline: 21333011]

49.   de BM, Sheeran P, Kok G, Hiemstra A, Prins JM, Hospers HJ, et al. Self-regulatory processes mediate the intention-behavior relation for adherence and exercise behaviors. Health Psychol 2012 Nov;31(6):695-703. [doi: 10.1037/a0027425] [Medline: 22390738]

50.   Albarracín D, Gillette JC, Earl AN, Glasman LR, Durantini MR, Ho M. A test of major assumptions about behavior change: a comprehensive look at the effects of passive and active HIV-prevention interventions since the beginning of the epidemic. Psychol Bull 2005 Nov;131(6):856-897 [FREE Full text] [doi: 10.1037/0033-2909.131.6.856] [Medline: 16351327]

51.   Webb TL, Sheeran P. Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. Psychol Bull 2006 Mar;132(2):249-268. [doi: 10.1037/0033-2909.132.2.249] [Medline: 16536643]

52.   Kremers SPJ, de Bruijn G, Visscher TLS, van Mechelen W, de Vries NK, Brug J. Environmental influences on energy balance-related behaviors: a dual-process view. Int J Behav Nutr Phys Act 2006;3:9 [FREE Full text] [doi: 10.1186/1479-5868-3-9] [Medline: 16700907]

53.   Olander EK, Fletcher H, Williams S, Atkinson L, Turner A, French DP. What are the most effective techniques in changing obese individuals' physical activity self-efficacy and behaviour: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2013;10:29 [FREE Full text] [doi: 10.1186/1479-5868-10-29] [Medline: 23452345]

54.   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]

55.   Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. Am J Community Psychol 2008 Jun;41(3-4):327-350. [doi: 10.1007/s10464-008-9165-0] [Medline: 18322790]

56.   de BM, Crutzen R, Peters GY. Everything should be as simple as possible, but this will still be complex: a reply to various commentaries on IPEBA. Health Psychol Rev 2015;9(1):38-41. [doi: 10.1080/17437199.2014.981833] [Medline: 25559162]

Abbreviations

BCT: behavior change technique
CALO-RE: Coventry, Aberdeen, and London—Refined taxonomy
CART: classification and regression trees
CMA: comprehensive meta-analysis
MoD: mode of delivery

XSL·FO
RenderX

properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

Original Paper

A Mobile Phone App Intervention Targeting Fruit and Vegetable Consumption: The Efficacy of Textual and Auditory Tailored Health Information Tested in a Randomized Controlled Trial

Sarah Pietertje Elbert[1*], PhD; Arie Dijkstra[1*], PhD; Anke Oenema[2*], PhD

[1]Department of Social Psychology, University of Groningen, Groningen, Netherlands

[2]CAPHRI School for Public Health and Primary Care, Department of Health Promotion, Maastricht University, Maastricht, Netherlands

[*]all authors contributed equally

Corresponding Author:
Sarah Pietertje Elbert, PhD
Department of Social Psychology
University of Groningen
Grote Kruisstraat 2/1
Groningen, 9712 TS
Netherlands
Phone: 31 503638033
Fax: 31 503636304
Email: s.p.elbert@rug.nl

Abstract

Background: Mobile phone apps are increasingly used to deliver health interventions, which provide the opportunity to present health information via different communication modes. However, scientific evidence regarding the effects of such health apps is scarce.

Objective: In a randomized controlled trial, we tested the efficacy of a 6-month intervention delivered via a mobile phone app that communicated either textual or auditory tailored health information aimed at stimulating fruit and vegetable intake. A control condition in which no health information was given was added. Perceived own health and health literacy were included as moderators to assess for which groups the interventions could possibly lead to health behavior change.

Methods: After downloading the mobile phone app, respondents were exposed monthly to either text-based or audio-based tailored health information and feedback over a period of 6 months via the mobile phone app. In addition, respondents in the control condition only completed the baseline and posttest measures. Within a community sample (online recruitment), self-reported fruit and vegetable intake at 6-month follow-up was our primary outcome measure.

Results: In total, 146 respondents (ranging from 40 to 58 per condition) completed the study (attrition rate 55%). A significant main effect of condition was found on fruit intake (P=.049, partial η^2=0.04). A higher fruit intake was found after exposure to the auditory information, especially in recipients with a poor perceived own health (P=.003, partial η^2=0.08). In addition, health literacy moderated the effect of condition on vegetable intake 6 months later (P<.001, partial η^2=.11). A higher vegetable intake was found for recipients with high health literacy after exposure to the textual or auditory intervention compared to the control condition (contrasts P=.07 and P=.004, respectively). In the case of relatively low health literacy, vegetable intake was the highest in the control condition (contrasts text control: P=.03; audio control: P=.04).

Conclusions: This study provides evidence-based insight into the effects of a mobile health app. The app seems to have the potential to change fruit and vegetable intake up to 6 months later, at least for specific groups. We found different effects for fruit and vegetable intake, respectively, suggesting that different underlying psychological mechanisms are associated with these specific behaviors. Based on our results, it seems worthwhile to investigate additional ways to increase fruit and vegetable intake in recipients with low health literacy.

ClinicalTrial: International Standard Randomized Controlled Trial Number (ISRCTN): 23466915; http://www.isrctn.com/ISRCTN23466915 (Archived by WebCite at http://www.webcitation.org/6hTtfSvaz)

XSL•FO
RenderX

Introduction

The number of mobile health apps is increasing rapidly. In 2013, more than 30,000 health apps were available [1], whereas more than 150,000 health-related apps were found only two years later [2]. In the broad context of mHealth, apps have been developed and used in health interventions, focusing on a wide array of behaviors and having different functions, such as the prevention and management of chronic and mental illnesses [3], and education about smoking, physical activity, or nutrition in an effort to stimulate behavior change in these domains [4,5]. However, scientific and evidence-based research with regard to the efficacy of these mobile health apps is lacking [3,6,7]. A recent review [6] shows that most published work on mobile apps used in health behavior change interventions are pilot studies [8] or describe the app in terms of content or acceptability [5,9]. Moreover, mobile phone apps are often offered as an additional tool to stimulate health behavior change next to an eHealth intervention, face-to-face counseling, or a virtual coach. To the best of our knowledge, only a few randomized controlled trials exist that focus on the effects of standalone mobile health apps [10,11], with only minimal follow-up periods ranging from 6 to 8 weeks. This means there is limited knowledge on the effectiveness of mobile phone apps in the process of health behavior change. In this study, the effects of auditory and textually tailored health information provided via a mobile phone app are tested in a randomized controlled trial.

Using mobile phone apps to deliver an intervention can have a variety of advantages. In addition to the increased availability and accessibility of mobile phones and the potential of reaching many people, it provides the opportunity to use interactive technological possibilities for persuasion that may support behavior change [12]. In particular, it enables the use of different communication modes (eg, text, video, and audio) and the use of computer tailoring to convey health information [13]. Mobile phones are, in general, already partly used for their MP3 function and mobile phone apps can be easily used to include and deliver auditory information, such as integrated within a health intervention. In addition, there is some evidence that at least audiovisual tailored messages can have advantages compared to text-based tailored messages [14,15] and at least one study suggests that audio-based information may be of added value in the stimulation of fruit and vegetable consumption [16]. Furthermore, tailoring can have beneficial effects in health interventions over providing nontailored information, for example, by increasing the relevance of the information [17-20]. The goal of this study is to investigate the efficacy of a mobile phone intervention that delivers tailored persuasive information as communicated via two different communication modes: text versus audio.

The mode of communication via which the persuasive health information is delivered might affect how the information is processed. For instance, compared to textually tailored information, interactive tailored information (either video- or audio-delivered) has been found to lead to greater attention [21,22] and is perceived as being more salient [23] and engaging [22]. In addition, in processing video- and audio-delivered communication, source considerations and peripheral cues or heuristics may play a more important role [24]. Furthermore, one study showed no significant differences between auditory and textual feedback on the recall of health-related information [25]. Other studies found mixed results between audiovisual and textual feedback [21,25]; audio only or audiovisual information was not always more effective than textual feedback. Thus, concerning the communication of health-related information, no explicit conclusion can be formulated with regard to the efficacy of a specific communication mode.

This intervention will apply and compare auditory and textual persuasive communication aimed at stimulating fruit and vegetable intake. A sufficient daily intake of fruit and vegetables contributes to the prevention of cardiovascular diseases and certain types of cancer [26]. However, more than 70% of the Dutch adult population does not meet the recommended minimum intake of fruit and this percentage is even higher for vegetable consumption [27]. These recommendations refer to a daily consumption of two pieces of fruit and 200 grams of vegetables for an adult population [28]. In addition, the average intake levels of fruit and vegetables seems to be decreasing over the years [27]. Moreover, similar intake patterns and trends are identified all over the world [29]. Thus, the stimulation of fruit and vegetable consumption remains a highly important health promotion topic.

To determine which communication mode can be used to deliver the tailored health information most effectively, it is important to test this in a randomized controlled trial. In this study, two research questions will be central. First, we aim to answer the question whether a tailored health intervention delivered via a mobile phone app is able to change fruit and vegetable intake in the advocated direction. Second, this study provides an exploratory test of the possible difference in efficacy between the more classic textual mode of communication (reading) and the auditory mode of communication (listening).

With regard to the first research question, it is expected that a tailored health intervention will be more effective compared to a control condition in which no health information is given. However, this difference may not be displayed in everyone, but only in a specific group of people. It is hypothesized that this will be especially the case in people who perceive a need to change their fruit and vegetable intake. It is reasoned that people who perceive their own health as relatively good have a lower need to change, whereas this need is higher for people who perceive their own health as relatively poor. The intervention might fit within the need for this latter group and, therefore, might be more beneficial for people who perceive their own health as relatively poor. Within the unimodel of persuasion [30], this could be described as a match between the persuasive

XSL•FO

RenderX

information and a premise held by a person (eg, "I might need this information because my health is not that good"), whereas this match might be lacking for people who perceive their own health as good in advance. Therefore, we will test the hypothesis that the intervention (either textual or auditory) will be more effective, especially for people who perceive their own health as relatively poor.

With regard to the second (exploratory) research question, it will be investigated whether the efficacy of the auditory intervention differs from the textual intervention. Again, this might not be the case for everyone. A relevant individual difference in this context is health literacy, defined as "the degree to which individuals can obtain, process, and understand the basic health information and services they need to make appropriate health decisions" [31]. Furthermore, health literacy is found to be related to level of education, cognitive and social skills, language and cultural barriers, and motivation [31,32], and low health literacy is associated with poorer health outcomes as well [33,34]. It seems worthwhile to consider the communication modality in combination with the construct of health literacy [34]. For instance, it is recommended to explore the use of auditory information because this might be especially

beneficial for people with low health literacy [32]. Therefore, it is expected that people with low health literacy may benefit from health information communicated via the auditory mode, whereas no specific differences are expected for people with high health literacy.

In summary, we aim to test the efficacy of two different fruit and vegetable promotion interventions delivered via a mobile phone app that communicates persuasive health information via an auditory or textual mode. The efficacy of the auditory and textual intervention will be compared with a control condition in which no intervention is present, and the textual and auditory interventions will be compared to each other. The content of the intervention is tailored to relevant characteristics of the individual: Feedback on the participants' perceived fruit and vegetable consumption and personalized recommendations regarding the individual barriers to eating sufficient fruit and vegetables are included. Other evidence-based behavior change strategies [35] applied in this intervention to assist behavior change are listed in Table 1. The dependent variables are self-reported fruit intake and self-reported vegetable intake at 6-month follow-up as assessed with a detailed and validated frequency questionnaire [36].

Table 1. Overview of behavior change techniques applied in the intervention.

Behavior change technique	Example
Tailored working mechanisms	
Provide feedback on performance	"You indicate that you eat sufficient fruit and vegetables, that is very positive" (baseline message)
Adaptation of the content	Provide information on how to overcome personal barriers; the inclusion of information about weight management based on dieting status (baseline message)
Preference tailoring	Use of preferred conversational form throughout the intervention
Testimonial matching	Exposure to a testimonial of the same gender as the respondent (follow-up moments)
General working mechanisms	
Provide information about behavior-health link	Provide general health risk information: "Eating sufficient fruit and vegetables contributes to good health" (baseline message)
Provide information on consequences	Provide positive outcomes of performing the behavior: "The vitamins, minerals, and fibers in fruit and vegetables affect your health in several ways" (low blood pressure, improved physical stamina, and decreased risk for diseases; baseline message)
Barrier identification	"Different aspects can play a role in eating insufficient fruit and/or vegetables. Of the following reasons, can you list maximally two aspects that apply to you?" (assessed at baseline)
Use of follow-up prompts	Monthly follow-up moments are created to encourage respondents to revisit the app
Provide opportunities for social comparison	The use of testimonials (follow-up moments)
Prompt specific goal setting	Respondents are encouraged to create own implementation intentions (general app content; menu button "action plan")
Repetition	Respondents can be exposed to the tailored message multiple times (general app content; menu button "my advice")

Methods

Recruitment

Participants were recruited in October and November 2013 to join a mobile phone intervention study aimed at stimulating

fruit and vegetable intake. Those interested were eligible for participation if they were 16 years or older, lived in the Netherlands, and owned an Android device (mobile phone or tablet, Android version 2.2 or later) with an installed version of Adobe Air (if necessary, they were automatically directed to

Google Play to install it safely). In collaboration with the programmers (affiliated with the University of Groningen), we decided to focus solely on the Android operating system because market research showed that the majority of Dutch mobile phone users owned an Android device [37]. In addition, the recruitment invitation was specifically aimed at people who had not (yet) succeeded in consuming two pieces of fruit and 200 grams of vegetables on a daily basis. After the 2-month recruitment period, interested people who signed up could not participate anymore.

Participants were recruited via several advertising campaigns published on newspaper and (health) magazine websites, on the local university website, in the online newsletter of the Netherlands Nutrition Centre, and via social networking websites. In addition, a local newspaper focused on the topic of fruit and vegetable consumption and referred to our mobile phone app, the "Fruit and Vegetables hAPP" (the addition of the "h" to the word "app" in Dutch means "snack" or "bite"). All advertisements briefly mentioned the content and 6-month duration of the study and provided a link to the mobile phone app in Google Play where respondents could find more information and download it. Respondents were not informed about the existence of different research conditions. They had a chance of winning different prizes (two Android tablet computers, 10 books with vegetable recipes, and 20 €10 coupons) after completing the pretest and posttest measurements. The study was approved by the ethical committee of the faculty of Behavioral and Social Sciences for conducting human participant research at the University of Groningen (no: 13012-N; trial registration: ISRCTN23466915).

We estimated the number of respondents to be included. This intervention and the comparison between auditory and textual persuasion were novel; therefore, it was difficult to predict what could be expected. We aimed to find medium effects for the comparison between one of the interventions and the control condition for the intervention to be of practical relevance. This meant at least 64 respondents needed to be included in each condition at posttest (P=.05, power=.80) [38].

Research Design and Procedure

This study was a pretest-posttest randomized controlled trial with two experimental conditions (text-based tailored health information and audio-based tailored health information) and a control condition in which respondents completed only the baseline measurements and posttest measurements at the 6-month follow-up. Those interested could download the mobile phone app in Google Play and sign up for the research via the app itself. This was done by creating a personal account with an email address and password, which was necessary to combine the data of the different measurements. In addition, the sign-up procedure consisted of questions on gender, first name, and preferred conversation form (in Dutch, a formal and polite

conversation form [*u, uw*] and a more informal conversation form [*jij, jou*] can be distinguished based on certain display rules). This information was used for tailoring purposes throughout the assessment and intervention. Next, respondents were presented with an informed consent form that stated the procedure, duration, and confidentiality of the research. In addition, it was mentioned that participation would be used for research purposes. After giving informed consent in the mobile phone app (with a checkbox), respondents were automatically assigned to one of the three conditions (sequentially in order of registration). All assessment questions (one question per screen) and the tailored health information were delivered via the mobile phone app. Figure 1 represents an overview of the design and the different elements of the study, which will be described subsequently.

Respondents in all conditions were asked to complete the pretest measures that consisted of baseline measures and questions for tailoring purposes. Respondents in the text-based and audio-based health information condition were then exposed to a tailored message (on the basis of decision rules) and additional evaluation measures. In total, this first contact took 20 minutes on average after which these respondents had access to the general mobile phone app content. Respondents in the control condition were only exposed to a message screen addressing the end of the baseline questionnaire. They were thanked for their participation and it was explicitly mentioned that they could expect another questionnaire 6 months later. They did not have access to the content of the mobile phone app.

Those respondents who did not complete the baseline questionnaire within 1 month were reminded by email to fill out the questionnaire. Further reminders were sent monthly; respondents who did not complete it during the research period were excluded from the study and informed about this by email.

In the months between the pretest and posttest assessments, respondents in the text-based and audio-based health information condition received monthly email invitations (with a maximum of four reminders during the month). They were asked to visit the mobile phone app to complete follow-up tailoring measures (identical for each month) and were exposed to newly added (either textual or auditory) tailored health information based on their input. Finally, at 6-month follow-up, all respondents were sent an email invitation to fill out the posttest measures in the mobile phone app (again with a maximum of four reminders during the month). Respondents who indicated they were interested in receiving more information were debriefed via email when the 6-month posttest had been completed. For ethical reasons, respondents could notify the researchers during the trial when they were not interested in participating anymore. After this notification, they did not receive monthly email invitations anymore, but only a final invitation to complete the posttest measures.

Figure 1. The design of the mobile phone health intervention to stimulate fruit and vegetable intake.
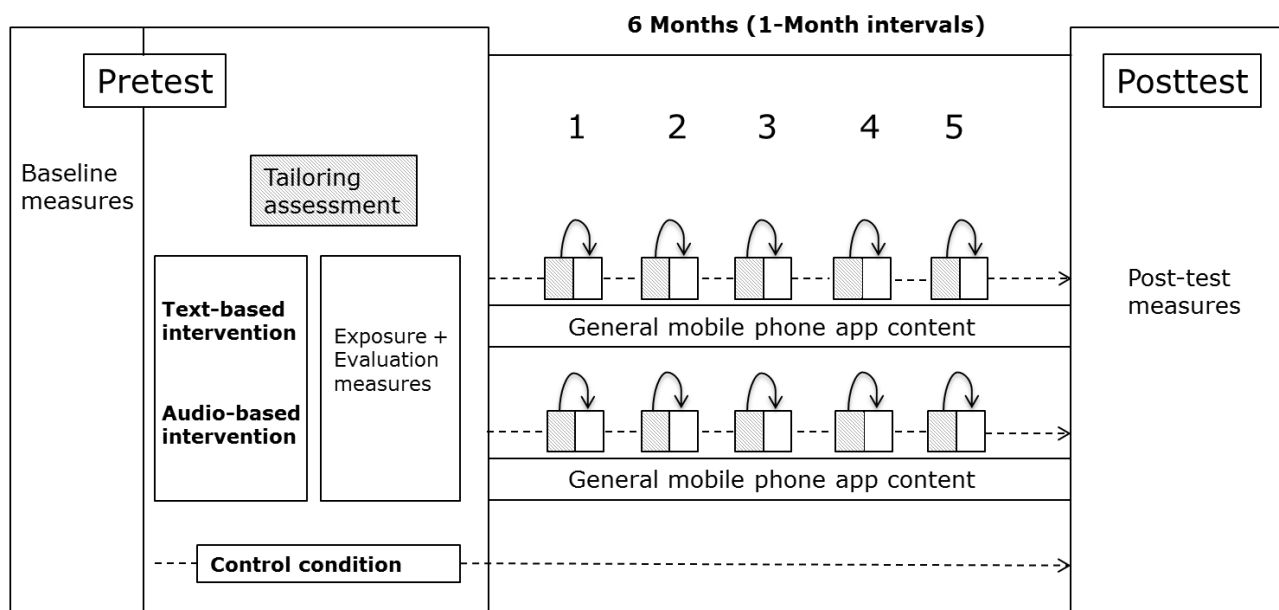
Intervention

This intervention was developed in the framework of Intervention Mapping [39-41]. In addition, behavior change techniques [35] were applied and the intervention was based on several sociocognitive determinants known to predict fruit and vegetable intake [42]. The core determinants included outcome expectations of eating sufficient fruit and vegetables and self-efficacy with regard to being able to eat sufficient fruit and vegetables. More specifically, the intervention focused on increasing positive outcome expectations and restructuring negative outcome expectations with regard to eating sufficient fruit and vegetables, and increasing self-efficacy. Different methods were used to address these factors. For instance, to target self-efficacy, we identified the respondents' experienced barriers to eating sufficient fruit and vegetables and provided relevant information to cope with these barriers [35].

The intervention consisted of one main moment of exposure to the tailored information at baseline and five follow-up moments with exposure to smaller components of tailored information. After exposure to the main tailored message right after completing the pretest measurements, respondents had access to additional functions of the mobile phone app throughout the 6 months. These functions were presented as seven main menu buttons (a screenshot is presented in Figure 2), including a button that invited respondents to formulate a personal action plan while making use of "if...then" formulations

(implementation intentions; Figure 2), a button with an alphabetical list of fruit and vegetables (Figure 2), and fruit and vegetable recipes. Four extra recipes were uploaded to the mobile phone app every month. In addition, one button included the most recent tailored message so it was possible to read or listen to it again (Figure 2).

During the five follow-up moments between the pretest and posttest assessments, respondents were exposed to a new, short tailored message each month that approached the topic of fruit and vegetable intake from distinct perspectives. Every month, the content was related to a general topic: the effect on well-being, the availability of fruit and vegetables, fruit and vegetables as a basic physical need, the lowered risk for chronic diseases, fruit and vegetable intake as a part of a healthy lifestyle, and objections people can have regarding fruit and vegetable intake, respectively. Additionally, a unique testimonial was included each month for sharing experiences on fruit and vegetable intake with the respondents. Testimonials are constructed stories in which successful personal experiences are shared to "directly or indirectly encourage the audience" to perform the behavior themselves [43]. For example, a physician elaborated on the relevance of a healthy diet and a nonexpert (without a specified occupation) expressed the experienced benefits of fruit and vegetables in the long term. To ensure that respondents were exposed to the information from different perspectives, the follow-up content was replaced in the mobile phone app at all five follow-up moments.

Figure 2. Screenshots of the app content: the main menu, starting an action plan, an alphabetical list of fruit and vegetables, and a tailored auditory advice. Three screenshots are translated from Dutch.
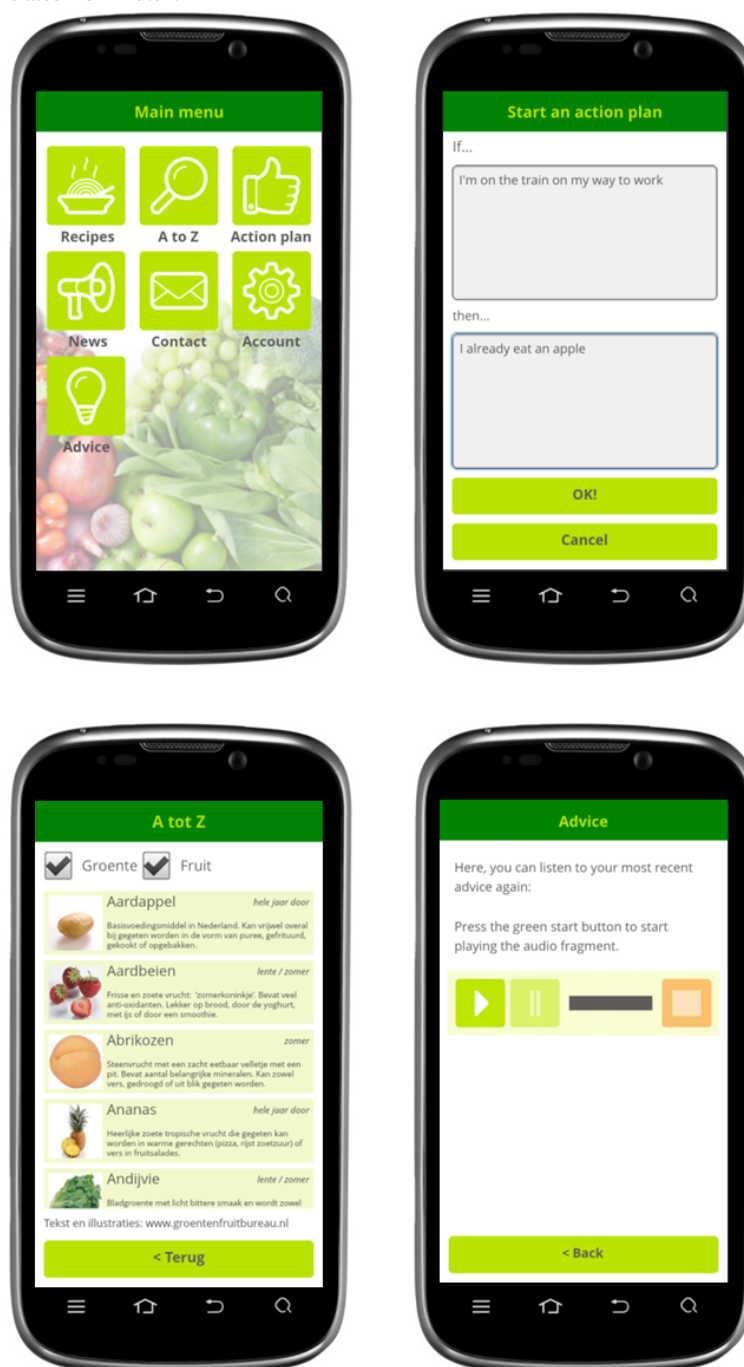
The Main Tailored Message

At baseline, a number of tailoring questions were included to partly determine the content of the feedback. Firstly, respondents could indicate with two questions whether their fruit and vegetable consumption was sufficient or not (perceived [subjective] consumption of fruit and vegetables, respectively), according to the recommendations. Participants could answer these questions with "Yes, I do meet this guideline;" "No, I probably do not always meet this guideline;" and "No, I do not meet this guideline." The second category was added to prevent people from overreporting their fruit and vegetable intake; the latter two answering options both reflected "not meeting the guideline."

Based on the answers given, respondents could select one or two individual barriers to eating sufficient fruit and/or vegetables from a predefined list. Respondents who indicated that they already met the guideline for eating sufficient fruit and vegetable consumption were asked to think of barriers in a future period in which they possibly would eat less fruit or/and vegetables. Examples of barriers included in the list were "I don"t like the taste of fruit" or "It takes a lot of time and effort to prepare vegetables" [44,45].

Another tailoring question concerned health value was assessed with one item ("How important is health to you?"). Participants could indicate whether they believed health was 1=not the most important to them (eg, "It is important to me, but not the most

important aspect in life") or 2=most important to them. In addition, they could indicate the perceived difference between themselves and their ideal and ought self, respectively ("In general, how large is the difference between who you actually are and who you prefer to be / who you should be?"). A scale was created by subtracting the score on the second item from the score on the first item; answering options ranged from 1=very small to 7=very large. The combination of health value and self-discrepancy determined whether the baseline message focused on the positive outcomes of sufficient fruit and vegetable consumption or on the negative outcomes of insufficient fruit and vegetable consumption [46]. Finally, it was assessed whether respondents were frequent dieters (answering options: 1=never, 2=sometimes, 3=regularly, 4=often) based on Lowe and Timko [47], and whether they had a partner relationship or not. These answers were used to decide whether or not to include information on the outcomes of sufficient/insufficient fruit and vegetable consumption related to weight management and appearance benefits, respectively.

Based on the answers on the previously mentioned tailoring questions, the baseline message consisted of a short general introduction providing information on the behavior-health link, feedback on their own fruit and vegetable consumption, and adapted information on the outcomes of sufficient/insufficient fruit and vegetable consumption with possibly information about appearance benefits and weight management. Then, feedback regarding one or two assessed individual barriers to perform the behavior with personal recommendations and formulation of relevant individual implementation intentions [48-50] were included. Throughout the mobile phone app, the preferred conversation form was used consistently. Transition sentences and closing sentences were created to ensure that the composed message was perceived as one fluent message.

The Follow-Up Tailored Messages

During all follow-up moments, respondents had to answer a maximum of four tailoring questions. First, the two questions on the current self-perceived fruit and vegetable intake were assessed again. The recommendations were included and respondents could indicate whether they met this guideline in the previous two weeks (answering options: yes/almost/no; the latter two answering options were considered as "not meeting the guideline"). In case of insufficient self-perceived fruit and/or vegetable consumption, we additionally asked whether the respondent had the intention to increase fruit and/or vegetable consumption in the following two weeks (answering options: no/a little/yes; the first two options reflected "no intention").

Based on the given answers, respondents were then exposed to a short (textual or auditory) feedback message that addressed a certain theme at each follow-up moment. After this, the (textual or auditory) testimonial was included that matched the respondent's own gender, except when it was one of the three testimonials that were only recorded with either a male or a female voice. This was decided in line with general expectations; for example, a dietician was only represented by a female voice. Respondents who already perceived their fruit and vegetable consumption as sufficient were not exposed to the thematic

information, but only to a short encouraging message and testimonial.

Mode of Delivery

The text-based and audio-based interventions varied only in their mode of delivery. The content information was partly composed in collaboration with the Netherlands Nutrition Centre and the auditory elements were developed in collaboration with a professional recording studio. An experienced female actor was selected for recording the baseline and follow-up feedback messages. She had a gender-congruent voice (feminine and high-pitched) and neutral sound without specific cultural or disturbing elements. After recording and arrangement sessions, the tailored audio files (233 files for the pretest and 114 for the follow-up moments, ranging from a single sentence to a text of 200 words) were mastered in 96 kHz 24 bit and converted to mono MP3 format (64 kbps) to use in the mobile phone app. Because it was important that the audio files of different parts could be arranged into one fluent message (without experiencing obvious transitions between parts), it was ensured that all recordings had a similar "tone of voice." Natural pauses between sentences lasted approximately 1 second; after every part, a 1-second pause was created as well to create as natural transition as possible. After a first evaluation round, the recording studio made some improvements; once the first author approved this, the audio files were uploaded in the intervention system by the programmers. Before listening to the baseline message, respondents in the audio-based information condition were presented with an instructive recording on volume regulation. They could adjust the mobile phone volume while listening to ascertain that it was sufficient and convenient. On the next screen, they could listen to the tailored health message. The complete message at baseline consisted on average of approximately 900 words (approximately 5 minutes for the auditory recording), roughly varying between 600 words (approximately 3 minutes) and 1200 words (approximately 6 minutes). In addition, the shorter tailored messages at follow-up consisted of 180 words on average and lasted 1:10 minutes on average for the auditory recording.

Contrary to the other auditory content, the testimonials within the follow-up moments were recorded with nonprofessional voices. The first author gave instructions and the testimonials were recorded in an office environment with a headphone microphone with the Praat software program [51] and arranged with the Audacity software program [52]. The recordings were send to the professional recording studio to make sure that the testimonials had the same quality and default volume as the remaining auditory content, again to ensure that it could be composed together. In total, 11 expert and nonexpert testimonials were created, among which four were recorded twice (ie, with a male and female voice). Three testimonials were only recorded for a man or a woman. On average, the testimonials consisted of 244 words and lasted 73 seconds. In total, an average follow-up moment lasted 2 to 3 minutes.

Multimedia Appendix 1 contains video and audio material of the mobile phone app content to provide insight into the registration procedure, baseline assessment, and exposure to the information at baseline and follow-up.

Measurements

Baseline Measurement

The following sociodemographic variables were assessed: age, cultural background, and highest level of completed education. This latter item was dichotomized into low (primary education, lower general secondary education, intermediate vocational education) and high level of completed education (higher general secondary education, higher vocational education, university level). Then, health-related questions were asked. The participants' perceived own health [53] was indicated on a 6-point scale ranging from ("my health is...") 1=very good to 6=very bad. This item was recoded and high scores corresponded with good perceived health (mean 4.86, SD 0.72). Self-reported height and weight were assessed to calculate body mass index (BMI) and we assessed whether respondents had a chronic disease or dyslexia. In addition, two items assessed perceived difficulty of eating sufficient fruit and sufficient vegetables as a measure of self-efficacy: "How difficult is it for you to eat sufficient fruit (vegetables)?" Both items could be answered on 5-point scales (1=not difficult at all, 2=not difficult, 3=neutral, 4=difficult, 5=very difficult). A composite measurement was created (r=.37, P<.001; mean 2.47, SD 0.95).

Then, we assessed the self-reported fruit and vegetable consumption in the previous month with a detailed and validated food-frequency questionnaire [36]. Respondents were asked how often, on average, per week they ate or drank products from several fruit and vegetable categories during the previous month. The answer options ranged from 0=never or less than 1 day a week, 1=1 day a week, to 7=every day. Next, they were asked to indicate the amount of intake per category of fruit or vegetables in terms of pieces of fruit and servings of vegetables, with the answering options ranging from 0=none, 1=1 piece, to 5=5 or more pieces. The main categories were cooked vegetables, raw vegetables/salad, fruit and vegetable juice, tangerines, oranges/grapefruits/lemons, apples/pears, bananas, other fruit, and apple sauce. The category "fruit and vegetable juice" was excluded because we would be unable to distinguish between fruit and vegetables. The number of days per week and the vegetable portions were multiplied for the first two categories and added to create a composite index of mean weekly vegetable intake for the previous month. The mean number of days per week and the fruit portions were multiplied for the remaining six categories and added to create a composite index of weekly fruit intake for the previous month.

Finally, we assessed respondents' health literacy with three statements that could be answered on 5-point scales ranging from 1=strongly disagree to 5=strongly agree. The three items were "I think it is easy to understand...information about health and lifestyle" / "...health information given by a physician, for example about a disease or treatment" / "...information about the effects of healthy nutrition" (α=.85, mean 4.23, SD 0.77).

Process evaluation questions were included immediately after respondents in one of the two experimental conditions were exposed to the tailored health information. These items were included to assess self-reported exposure ("Did you read/listen to the fragment?" answering options ranging from 1=yes, completely to 5=no, not at all) and potential distracting elements while reading or listening ("Was the reading or listening possibly disrupted, for example, by other people, hard sounds, music, or other distracting elements?" with answering options "yes" and "no"). In addition, the novelty and usefulness of the information were assessed with two statements ("The information was new to me / useful for me") that could be answered on a 5-point scale ranging from 1=strongly disagree to 5=strongly agree. Finally, a general evaluation question was included ("How would you rate the intervention?"). This item could be answered on a 7-point scale ranging from 1=very negative to 7=very positive.

Posttest Measurement

At 6-month follow-up, fruit and vegetable intake was again assessed with the same questionnaire as at baseline [36]. Again, two composite measures for fruit and vegetable consumption, respectively, were created. In addition, it was assessed how often respondents searched for information about health and fruit and vegetables besides the information in the mobile phone app (mean 3.03, SD 1.03) and to what extent they spoke to others about the topic in the past 6 months (mean 2.71, SD 1.08). Both items could be answered on 5-point scales with answering options ranging from 1=never to 5=often. Finally, seven questions were added to evaluate the information and mobile phone app as a whole for a range of measures, such as personal applicability, novelty, credibility, the extent to which it is perceived as intense, usefulness, comprehensibility, and visual attractiveness. These questions could be answered on 5-point scales, with answering options ranging from 1=strongly disagree to 5=strongly agree. In addition, one item provided recipients with the opportunity to give qualitative feedback.

Analyses

First, univariate analyses (ANOVA, chi-square) were conducted to analyze whether the respondents in the conditions differed on relevant pretest measures and to see whether respondents who dropped out after baseline (184/329, 55.9%) significantly differed from the respondents who completed both measurements. Second, ANCOVAs were conducted for fruit and vegetable intake separately, while controlling for self-reported fruit or vegetable intake at baseline, age, and highest completed education (because these two latter variables had a large variance within our community sample). After testing the main effects of condition, two moderators (perceived own health status and health literacy) were tested on fruit and vegetable intake to assess whether the effects of condition were similar in specific groups of respondents. The same covariates were included. To examine the meaning of the moderation effects, simple main analyses were conducted at two different levels (low/high) of the moderator. The complete dataset was then used to model respondents as scoring high or low by adding and subtracting 1 SD to the standardized means, respectively [54]. Post hoc contrasts were inspected to investigate the difference between one of the interventions and the control condition, and to explore differences between both interventions. In addition, in the intention-to-treat analyses, the posttest (T2) fruit and vegetable intake of respondents who did not complete the study were considered to be equal to their reported fruit and vegetable intake at pretest. With these data, we again conducted the analyses on fruit consumption and vegetable consumption.

Finally, two categories of process measures were inspected referring to exposure to the intervention and a general evaluation of the main tailored message at baseline and the complete intervention. Age and education were now applied as covariates.

Results

Sample Characteristics

In total, 342 respondents registered for the study and started the pretest measurement and 96.5% completed it (330/342). Of these 330 respondents, 147 respondents (44.5%) completed the final questionnaire at 6-month follow-up as well. One respondent was excluded because he or she reported a fruit intolerance. The final sample for the analyses consisted of 146 respondents of whom 73.3% were females (n=107), 71.2% were highly educated (n=104), varied in age from 16 to 71 years (mean 41.4, SD 14.6), and with a mean BMI of 25.2 (SD 5.5). Recipients with a lower education reported a lower health literacy (mean 4.01, SD 0.77) compared to recipients with a higher education (mean 4.32, SD 0.76, P=.03). Thirty of 330 respondents (9.1%) were accidentally exposed to the same information during the first two follow-up moments (for technological reasons), instead of being exposed to different content information (referring to 20 respondents in the sample that actually completed the study, 13.7%).

The composite index of fruit and vegetable intake at pretest was treated as an indication of self-reported fruit and vegetable intake. The mean fruit intake was sufficient (14 portions considered sufficient; scale ranging from 0 to 56; mean 14.04, SD 10.63), whereas the mean vegetable intake was insufficient (28 portions considered sufficient; scale ranging from 6 to 70; mean 25.44, SD 11.37). If one of the two questions was answered with zero ("never or less than 1 day a week" or "no servings"), the total intake for that specific category was automatically set at zero as well (this was also when the pattern was not filled in consistently). This conservative approach means that the fruit and vegetable intake might be lower than in reality. Condition did not affect whether a question on fruit or vegetable intake was filled in consistently or not at baseline (P=.66) or posttest (P=.64).

Based on the answers given, 19.2% (28/146) of the respondents were classified as consuming an insufficient amount of vegetables (but a sufficient amount of fruit), 16.4% (24/146) were classified as consuming an insufficient amount of fruit (but a sufficient amount of vegetables), 48.6% (71/146) were classified as consuming both insufficient amounts of fruit and vegetables, and 15.8% (23/146) were classified as consuming sufficient amounts of both fruit and vegetables. At posttest, the scores on the composite index of fruit and vegetable intake were somewhat higher (fruit: scale ranging from 1 to 56.5, mean 14.93, SD 9.27; vegetables: scale ranging from 6 to 70, mean 27.47, SD 11.81). The respondents were distributed over the conditions as follows: textual health information (n=48), auditory health information (n=40), control condition (n=58). Figure 3 represents a flowchart in which the dropout rates per condition are depicted.

Figure 3. Flowchart of number of participants allocated per condition.
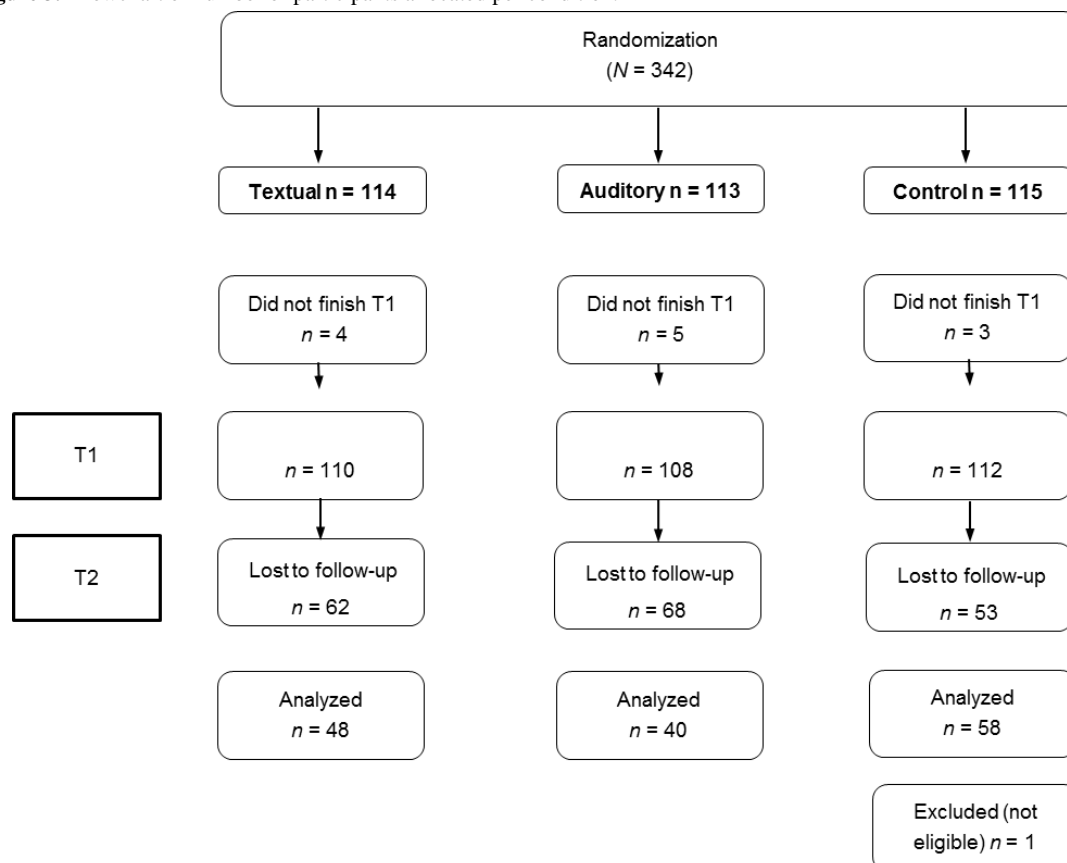
Randomization Check and Attrition Analyses

First, no significant differences between conditions were found regarding our set of 18 demographic and health-related baseline variables: gender ($P=.53$), age ($P=.11$), highest completed education ($P=.35$), cultural background ($P=.75$), dieting status ($P=.09$), relationship status ($P=.41$), the extent to which health is valued ($P=.25$), discrepancy between ought and ideal self ($P=.54$), having dyslexia ($P=.94$), perceived own health status ($P=.78$), having a chronic disease ($P=.84$), BMI ($P=.30$), self-efficacy ($P=.72$), self-reported fruit consumption ($P=.79$), self-reported vegetable consumption ($P=.16$), perceived (subjective) fruit and vegetable consumption ($P=.86$ and $P=.37$, respectively), and health literacy ($P=.07$).

In addition, the respondents who dropped out after baseline and the respondents who completed both measurements did not significantly differ on the pretest measures as mentioned in the previous paragraph (all $P>.12$). Furthermore, condition did not significantly affect whether respondents dropped out during the trial ($P=.08$). However, respondents who completed the study and who received either the auditory or textual feedback reported a significantly higher extent of being exposed to the information ($P<.001$) and they had a slightly more positive general impression of the pretest intervention content and measures compared to those who dropped out ($P=.06$). Finally, when exposed to the auditory feedback, respondents who dropped out reported being distracted while listening to the baseline intervention content more often compared to those who completed the whole study ($P=.03$). No significant differences were found on the extent to which the information was perceived as new or useful.

Effects on Fruit and Vegetable Consumption

Main Effects

We assessed whether condition affected the self-reported intake of fruit and vegetables 6 months after baseline. With regard to fruit intake, a significant main effect was found ($F_{2,140}=3.08$, $P=.049$, partial $\eta^2=0.04$) with the following estimated means: text (mean 13.5, SE 1.0), audio (mean 17.1, SE 1.2), and control (mean 14.3, SE 0.9). The difference between text and audio was significant ($P=.02$), but it did not reach statistical significance between audio and control ($P=.06$). The difference between text and control was not significant ($P=.53$). No significant main effect was found on vegetable intake ($F_{2,140}=0.01$, $P=.99$, partial $\eta^2=0.00$).

For fruit consumption, the raw means at pretest and posttest were inspected to gain more insight into the actual differences per condition. The means remained quite similar in the textual feedback condition (pretest vs posttest: mean 14.8, SD 11.1 vs mean 14.2, SD 6.9) and in the control condition (pretest vs posttest: mean 13.4, SD 10.4 vs mean 13.8, SD 9.4). Thus, only small differences were observed here (mean –0.6 pieces and 0.4 pieces per week, respectively). In the auditory feedback condition, the fruit intake was most strongly increased (3.3 pieces; pretest vs posttest: mean 14.2, SD 10.6 vs mean 17.5, SD 11.1).

Moderation Effects

A significant interaction was found between condition and perceived own health status on fruit intake ($F_{2,137}=4.24$, $P=.02$, partial $\eta^2=0.06$), but not on vegetable intake ($F_{2,137}=0.15$, $P=.86$, partial $\eta^2=0.00$). Figure 4 displays the mean fruit consumption for respondents with poor and good perceived own health.

In case of poor perceived own health status, condition did significantly affect fruit consumption at 6-month follow-up ($F_{2,137}=6.05$, $P=.003$, partial $\eta^2=0.08$). The mean scores were as follows: text (mean 14.2), audio (mean 20.5), and control (mean 13.2). Post hoc contrasts showed that the intake of fruit was significantly higher after listening to the information compared to the other two conditions (text: $P=.006$; control: $P=.001$). In case of good perceived own health status, condition had no significant effect ($F_{2,137}=1.15$, $P=.32$, partial $\eta^2=0.02$). The mean scores were as follows: text (mean 13.3), audio (mean 13.8), and control (mean 15.9). No significant contrasts were found.

Second, health literacy as assessed at pretest was tested as a moderator. No significant interaction with condition was found on fruit intake ($F_{2,137}=0.25$, $P=.78$, partial $\eta^2=0.00$). However, we found a significant interaction on vegetable intake ($F_{2,137}=8.42$, $P<.001$, partial $\eta^2=0.11$). Figure 5 displays the mean vegetable intake in the conditions for people with relatively low and high health literacy.

In case of low health literacy, condition significantly affected vegetable consumption at 6-month follow-up ($F_{2,137}=3.62$, $P=.03$, partial $\eta^2=0.05$). The mean scores were as follows: text (mean 21.3), audio (mean 23.1), and control (mean 27.9). Post hoc contrasts showed that the intake of vegetables in this group was significantly higher in the control condition compared to the two interventions (text: $P=.03$; audio: $P=.04$). In case of high health literacy, condition significantly affected vegetable consumption at 6-month follow-up as well ($F_{2,137}=4.53$, $P=.01$, partial $\eta^2=0.06$). The mean scores were as follows: text (mean 30.1), audio (mean 33.5), and control (mean 25.6). In addition to the higher scores compared to respondents with low health literacy, post hoc contrasts showed that the intake of vegetables in this group was lower in the control condition compared to the textual intervention ($P=.07$) and the auditory intervention ($P=.004$).

The main effect of condition was not significant on a composite measure of fruit and vegetable consumption ($P=.34$) and neither was the interaction with perceived own health status ($P=.16$). However, the interaction with health literacy was ($F_{2,137}=4.39$, $P=.01$, partial $\eta^2=0.06$). The main effect of condition then became nonsignificant in respondents with low health literacy and only two of the four contrasts remained significant, showing that the control condition was most effective for low-literate respondents (compared to reading) and listening to the auditory information was most effective in high-literate respondents (compared to control).

We further analyzed the effects in selections of respondents for whom the health information could be especially relevant. With regard to the effects on fruit intake, respondents were selected whose self-reported fruit intake was found to be insufficient at pretest (and the vegetable intake was either insufficient or sufficient, n=95). The main effect on fruit consumption was not significant anymore ($F_{2,89}$=1.85, P=.16, partial η^2=0.04). The estimated means were lower showing a similar pattern: text (mean 11.2, SE 1.12), audio (mean 13.7, SE 1.25), and control (mean 10.9, SE 1.00). The contrasts were not significant anymore (all P>.07), but the pattern of findings remained similar. Again, this pattern was especially found for recipients with a poor perceived own health (moderation effect: $F_{2,86}$=2.46, P=.09, partial η^2=0.05; main effect: $F_{2,86}$=3.95, P=.02, partial

η^2=0.08), showing significant differences between control (mean 8.9) and audio (mean 16.3; P=.006), and text (mean 10.5) and audio (P=.04) for this group. No significant differences between the conditions were found for recipients who perceived the own health as relatively good. As for the whole sample, health literacy did not moderate the effect on fruit intake.

In addition, the effects on vegetable intake were analyzed in a selection of respondents who indicated eating insufficient vegetables at pretest (and either insufficient or sufficient fruit intake, n=99). As in the whole sample, no significant main effect or moderation of perceived own health was found; however, the moderation of health literacy was not significant anymore ($F_{2,90}$=0.07, P=.93, partial η^2=0.00).

Figure 4. The interaction between condition and perceived own health status on fruit consumption at 6-month follow-up controlled for age, highest completed education, and self-reported fruit intake at baseline.
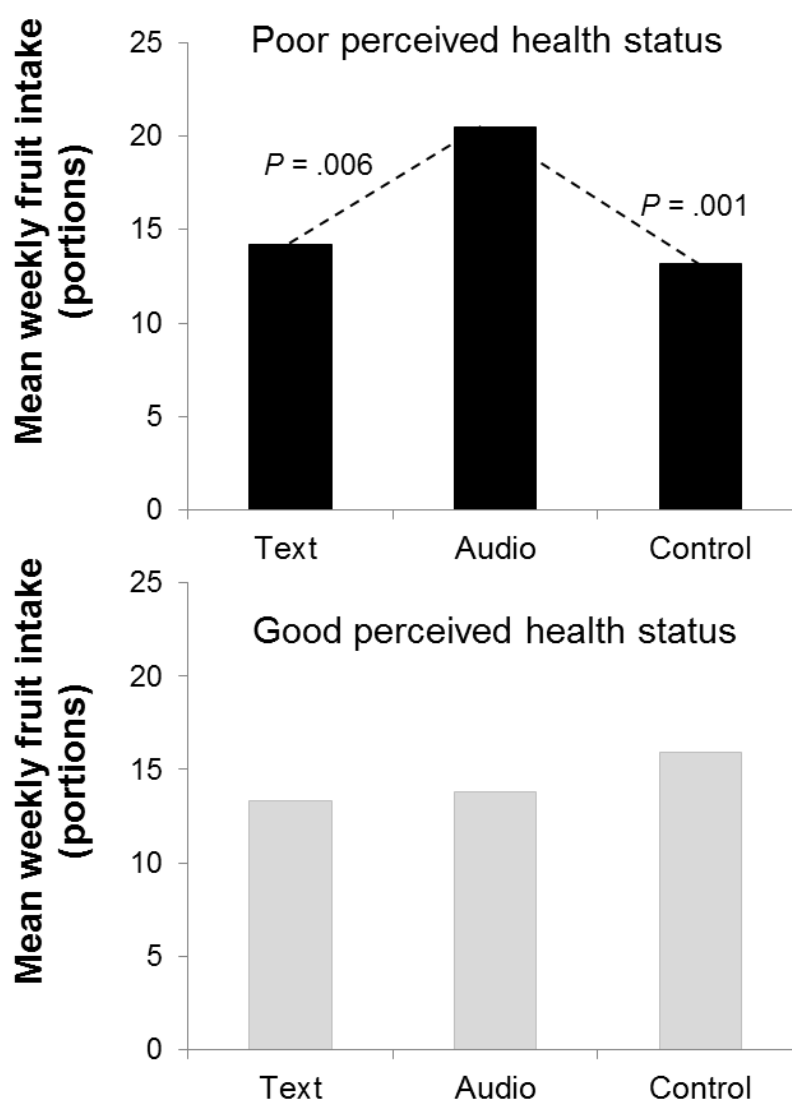
Figure 5. The interaction between condition and health literacy on vegetable consumption at 6-month follow-up controlled for age, highest completed education, and self-reported vegetable intake at baseline.
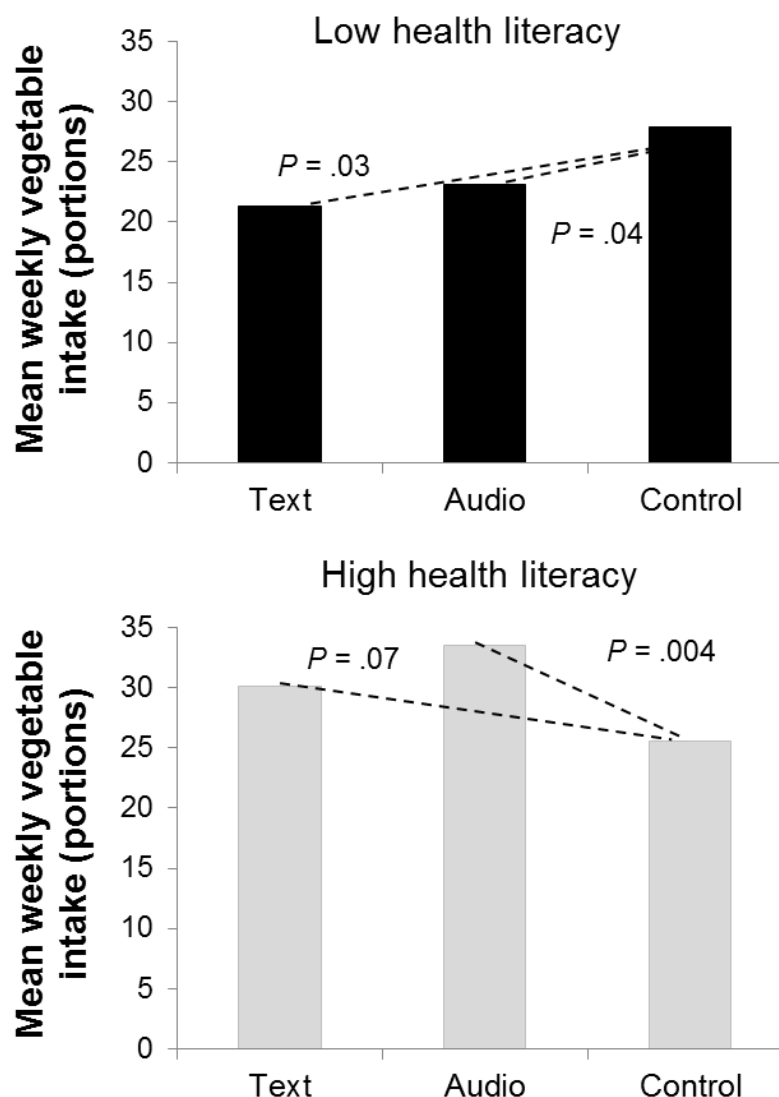
Intention-to-Treat Analyses

At T2, 183 of 330 (55.5%) respondents had dropped out despite reminders to fill in the follow-up measurement. In the intention-to-treat analysis (while assuming that the fruit and vegetable intake was equal to the assessed fruit and vegetable intake at pretest for the respondents who dropped out), we again conducted the analyses on fruit consumption and vegetable consumption. The main effect on fruit consumption remained significant ($F_{2,323}=3.18$, $P=.04$, partial $\eta^2=0.02$) with the following means: text (mean 13.5, SE 0.52), audio (mean 15.3, SE 0.52), and control (mean 14.3, SE 0.51). Now, the interaction between condition and perceived health status on fruit consumption was not significant anymore ($P=.33$). However, as expected, for respondents with a poor perceived health status, the effect of condition remained significant ($P=.04$, partial $\eta^2=0.02$) with a similar pattern of means and significant contrasts (audio vs text: $P=.02$; audio vs control: $P=.04$) compared to the original analyses. The interaction between condition and health literacy on vegetable consumption remained significant as well ($F_{2,320}=5.52$, $P=.004$, partial $\eta^2=0.03$). For both respondents with low and high health literacy, the effect of condition became marginally significant (all $P=.07$) with only the contrasts between audio-based health information and the control condition being significant (all $P=.02$). Overall, small(er) effect sizes were found.

Process Analyses

Effects on two categories of process variables were inspected, referring to the exposure to the intervention and a general evaluation of the main tailored message at baseline and the intervention in general. First, with regard to exposure, participants logged in a mean 7.6 times (SD 4.5) and, as expected, this was significantly more in one of the experimental conditions ($F_{2,143}=42.11$, $P<.001$, partial $\eta^2=0.37$, text: mean 10.1, SD 3.7; audio: mean 9.4, SD 4.2) compared to the control condition (mean 4.2, SD 3.0; contrasts text control and audio control; both $P<.001$,). No significant difference was found between the two interventions ($P=.35$).

XSL•FO
RenderX

On average, respondents completed 4.1 follow-up moments (SD 1.36). There was a difference between the two interventions: after reading, recipients completed slightly more follow-up moments (mean 4.35, SD 1.19) compared to the recipients who listened to the information (mean 3.83, SD 1.50; $F_{1,86}$=3.40, P=.07, partial η^2=0.04). In addition, 55 of 88 (63%) respondents in one of the experimental conditions completed all five follow-up moments and at least two follow-up moments were completed by 84 of 88 (95%) of the respondents. Those respondents who completed four or five follow-up moments were selected (text: n=39; audio: n=25) and compared to the control group (n=58) with regard to the set of 18 baseline measures. The groups did not differ significantly from one another regarding these variables (all P>.06).

At baseline, respondents who were in one of the intervention conditions were asked to indicate the extent to which they were exposed to the main tailored message and the extent to which they experienced potential distracting elements. Slightly more respondents reported being only partly exposed after listening ($F_{1,214}$=2.94, P=.09) and fewer respondents identified potential distracting elements after listening to the main tailored message compared to those who read the information (χ^2_1=2.7, P=.10).

Second, differences between conditions were found on perceived usefulness and the general evaluation of the main tailored message ($F_{1,214}$=12.27, P=.001, partial η^2=0.05 and $F_{1,214}$=12.10, P=.001, partial η^2=0.05, respectively): Respondents who *read* the baseline information experienced it as significantly more useful (mean 4.21, SE 0.08) and positive (mean 5.77, SE 0.09) compared to the respondents who *listened* to the baseline information (usefulness: mean 3.81, SE 0.08, general evaluation: mean 5.33, SE 0.09). No significant differences were found on the perceived novelty of the information (P=.29).

With regard to the evaluation of the mobile phone app and intervention content at the 6-month follow-up, respondents who were exposed to the audio-based content reported to have looked more often for additional information about health and fruit and vegetables mostly via Internet websites ($F_{2,141}$=3.00, P=.05, partial η^2=0.04) compared to respondents in the control condition (contrast P=.02). In addition, after exposure to the audio-based intervention respondents appeared to have talked more about the topic with other people ($F_{2,141}$=2.49, P=.09, partial η^2=0.03) compared to the control condition (contrast P=.03). When comparing both interventions, the feedback and the mobile phone app were experienced equally in terms of personal applicability, novelty, credibility, intensity, usefulness, comprehensibility, and visual attractiveness.

Discussion

This study addressed the efficacy of two tailored mobile phone interventions in a sample of people who were invited to participate especially when they perceived their own fruit and vegetable intake as insufficient. The efficacy of the interventions was compared to a control condition in which no tailored health information was provided, and an exploratory comparison was made between the text-based and audio-based tailored intervention. Besides testing this main effect, two relevant moderators—perceived own health status and health literacy—were included in this research.

Principal Results

It seems that the results for fruit consumption and vegetable consumption are different. The results on fruit consumption were supported by the similar findings of the intention-to-treat analysis and, although the effects were less strong in the selection of respondents, the pattern of findings was still present in respondents with a more objectively assessed fruit consumption that was insufficient (according to guidelines). The significant main effect showed that the audio-based intervention was more effective than both the text-based intervention and the control condition. The auditory mode of communication, but not the textual mode of communication, led to increased fruit consumption with a mean increase of three pieces of fruit a week. We did not expect specific differences between the efficacy of textual and auditory health information, but it is in line with previous studies on the potential efficacy of auditory information [16]. The auditory information may have led to more attention [21,22] or it may have been perceived as more rich and personal to the recipient [23], which is possibly translated into behavior change.

With regard to vegetable consumption, there was no significant main effect of condition. Instead, an interaction between condition and health literacy was found, which was supported by the intention-to-treat analysis, but the pattern of findings was not found in the selection of respondents with a more objectively assessed vegetable consumption that was insufficient. Yet, there was no difference between the auditory and textual health information: both the text-based and audio-based intervention led to higher vegetable consumption in respondents with high health literacy, whereas this was not the case for respondents with relatively low health literacy (within this highly educated sample). For them, both interventions led to a significant decrease in self-reported vegetable consumption at the 6-month follow-up compared to the control condition. It seems that the current mobile phone app was not helpful for people with relatively low health literacy, at least not in improving a complex behavior such as vegetable intake.

We found that respondents who perceived their own health as relatively poor reported higher fruit consumption after being exposed to the auditory health information. It was initially expected that the health information in general would be more relevant for recipients with a poor perceived own health because they might perceive the necessity to change and are willing to make more investments [43,55]. In other words, there is a match between the persuasive health information and a characteristic of the recipient [30]. Thus, it seems that recipients with poor perceived own health did benefit most from the rich and personal auditory information. Although speculating, this might be related to an optimal level of threat of the auditory information that was necessary to engage in behavior change or the promise that the threat will be lowered once the recipient engages in the behavior. It is important to address these underlying processes

in further research because the findings on the current process evaluation measures are unlikely to explain this pattern.

The moderator effect on vegetable intake shows that the intervention in general seemed to have worked especially in respondents with high health literacy. It can be that recipients with relatively low health literacy did not understand all content information or were not motivated to process the information [32,34] and therefore discarded the information in general. For recipients with high health literacy, it did not seem to matter how the information was communicated because they may have been open to the content information regardless of the mode of communication. It is important to unravel the different aspects of health literacy; for instance, education could have played a relevant role because recipients with a lower education level also reported lower health literacy in this study, but it is also possible that low health literacy is related to a defensive reaction to threatening health information.

Thus, we observed a main effect on fruit intake and, within a subsample of respondents, we could also find effects on vegetable intake. However, this finding on vegetable intake could not be replicated within a subsample of recipients who indicated consuming insufficient vegetables. This suggests that the findings on vegetable intake are less robust than on fruit intake. In addition, according to a conventional rule of thumb [56], small to medium effect sizes are found for fruit intake ranging between partial $\eta^2=0.04$ and partial $\eta^2=0.08$. It remains the question whether the absence of a main effect on vegetable intake was a matter of power because the recommended number of respondents per condition (n=64) was not reached. Yet, this effect might then have been too small to be relevant and the moderating effect of health literacy on vegetable intake showing contradictory results for recipients with low and high health literacy may indicate that it is indeed unlikely to find a main effect on vegetable intake.

The results suggest that not everybody benefitted equally from the intervention and that it can even adversely affect vegetable consumption. Thus, the tailored health information may have negative effects in subgroups of respondents when the information is communicated either via a textual or auditory channel. This increased our awareness on possible side effects of persuasive health communication. For future research, it seems worthwhile to investigate how individual characteristics can be assessed to optimize the practice of persuasive health communication and to increase knowledge on how "hard-to-reach" groups can benefit from it as well [32]. Possibilities may not only lie in providing persuasive tailored health information via another communication modality (ie, video tailoring), but also in the use of other interactive methods and elements (ie, the use of sensors, serious games, or avatars).

We did not expect specific differences for fruit and vegetable intake. It may be that these differences are found because fruit and vegetables are products with different qualities with regard to taste, preparation and culinary uses [57], and perceived ease of increasing consumption [55]. It may be that after a follow-up period of 6 months, the auditory intervention was only able to change fruit intake because this is a relatively less difficult behavior to change. In addition, interventions may differentially

lead to increased fruit and vegetable consumption [58]. For instance, the context in which the products are consumed might play a role: vegetables are more likely to be consumed in a social context (at dinner, with the rest of the family), whereas fruit is more likely to be consumed individually. Therefore, environmental interventions may be more effective for increasing vegetable intake. It seems a rational choice to assess fruit and vegetable consumption separately in future studies.

The effects on the process measures showed that there were no differences between the text-based and audio-based interventions with regard to exposure. However, more follow-up moments were completed by recipients who read the information compared to recipients who listened to the information. In addition, after the first contact, the baseline textual information was evaluated as more useful and positive compared to the auditory information. However, after 6 months the textual and auditory interventions were not evaluated significantly different on relevant measures, such as novelty, credibility, comprehensibility, and usefulness. Yet, recipients who were exposed to the auditory information searched more for health-related information and they discussed the topic more often with other people compared to the control condition. Further research is needed to explain these findings and to investigate how characteristics of the textual and auditory information may possibly have contributed to these differences.

Limitations

This research also has several limitations. First, a high percentage of respondents did not complete the entire study, which might have biased these findings. High attrition rates are common in Internet-based health behavior change interventions [59,60]. Although initially one could reason that an app might lead to lower attrition rates because people can participate in the study anywhere and at any time, it seems that for mobile phone apps it is still a challenge to keep respondents involved after their first visit. This might illustrate the quick and shallow relationship people have with the Internet and mobile phone apps in general ("app snacking"). Yet, it is found that respondents who dropped out from the research did not differ in relevant pretest measures compared to the respondents who completed the study.

Although it was not a significant result in this study, more respondents dropped out in one of the intervention conditions, which is a common finding in health intervention research [60]. Respondents in the intervention conditions were sent email reminders frequently and they were informed about the possibility to end their participation via email, whereas this was not the case for respondents in the control condition. These aspects of our research might additionally have contributed to differences in attrition between the conditions. A specific improvement may refer to sending reminders as mobile phone notifications versus email messages.

Secondly, we aimed to increase exposure by sending email prompts and providing regular updates of intervention content [59]. However, it is difficult to detect the actual exposure to the intervention content. People may report that they were fully exposed to the information, but still we do not know the quality of the exposure. In addition, we could not test intervention

components separately, which means we do not know specifically why respondents showed certain improvements. Thus, we were not able to determine the unique contribution of each component of the current mobile phone intervention and to make statements about the elements that affected fruit and vegetable intake specifically.

Thirdly, the sample was a selective community sample, which could have biased the results. Respondents were not necessarily a representation of the whole community because they were mostly highly educated with a Dutch nationality and had to use an Android device to be included in the research. Furthermore, in our recruitment, people were invited who did not eat sufficient fruit and vegetables, which is obviously a selection of respondents who might be interested in the topic of health and changing their health behavior. In addition, people tend to generally overestimate their fruit and vegetable intake with self-report assessment measures, as used in this study [61].

Conclusions

In our view, this app may be an effective channel to change fruit and vegetable intake, at least in certain groups of respondents. The development of the audio-based content was more costly and time-consuming compared to the text-based content, but it has shown to have beneficial effects on fruit consumption, or at least for some subgroups. The results showed

us that it is important to be aware of the possible side effects of psychological health interventions and to take into account individual differences when exposing respondents to threatening health information and personal feedback on fruit and vegetable intake.

A next step may be to optimize the mobile phone app. It is worthwhile to investigate possibilities to expose the subgroups to either one of the current interventions that was shown to be efficient, depending on the specific behavior one would like to change. Furthermore, tracking and sensor technologies can be added to use the mobile phone app as an intervention channel to its fullest potential [62], which means that recipients can keep track of their daily fruit and vegetable intake and may receive reminders to buy fruit and vegetables when they are in the supermarket. In addition, it would be worthwhile to ensure a higher level of interactivity between the recipient and the mobile phone app as an interactive information system.

To the best of our knowledge, this is a first test of the effects of communication modalities in an evidence-based tailored mobile phone app to stimulate fruit and vegetable intake. It provided us with new insights on the efficacy and processes involved, and we hope to inspire the testing and development of evidence-based mobile phone apps in the field of health education and promotion.

Conflicts of Interest

None declared.

Multimedia Appendix 1

A powerpoint presentation with additional video and audio material of the smartphone application content.

[PPTX File, 85MB - jmir_v18i6e147_app1.pptx]

Multimedia Appendix 2

CONSORT EHEALTH Checklist V1.6.

[PDF File (Adobe PDF File), 948KB - jmir_v18i6e147_app2.pdf]

References

1. Essany M. mHealthWatch. 2013 Apr 15. Mobile health care apps growing fast in number URL: http://mhealthwatch.com/mobile-health-care-apps-growing-fast-in-number-20052/ [accessed 2016-02-15] [WebCite Cache ID 6fJe1tsEE]

2. Misra S. iMedicalApps. 2015 Sep 17. New report finds more than 165,000 mobile health apps now available, takes close look at characteristics and use URL: http://www.imedicalapps.com/2015/09/ims-health-apps-report/ [WebCite Cache ID 6hZikREi8]

3. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. J Med Internet Res 2013;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]

4. Hebden L, Cook A, van der Ploeg HP, Allman-Farinelli M. Development of smartphone applications for nutrition and physical activity behavior change. JMIR Res Protoc 2012;1(2):e9 [FREE Full text] [doi: 10.2196/resprot.2205] [Medline: 23611892]

XSL•FO
RenderX

5. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res 2013;15(4):e86 [FREE Full text] [doi: 10.2196/jmir.2583] [Medline: 23598614]

6. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR Mhealth Uhealth 2015;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]

7. Wiederhold BK. Behavioral health apps abundant, but evidence-based research nearly nonexistent. Cyberpsychol Behav Soc Netw 2015 Jun;18(6):309-310. [doi: 10.1089/cyber.2015.29001.bkw] [Medline: 26075916]

8. Brindal E, Hendrie G, Freyne J, Coombe M, Berkovsky S, Noakes M. Design and pilot results of a mobile phone weight-loss application for women starting a meal replacement programme. J Telemed Telecare 2013 Mar 21:e1. [doi: 10.1177/1357633X13479702] [Medline: 23520213]

9. Cowan LT, Van Wagenen SA, Brown BA, Hedin RJ, Seino-Stephan Y, Hall PC, et al. Apps of steel: are exercise apps providing consumers with realistic expectations?: a content analysis of exercise apps for presence of behavior change theory. Health Educ Behav 2013 Apr;40(2):133-139. [doi: 10.1177/1090198112452126] [Medline: 22991048]

10. Glynn LG, Hayes PS, Casey M, Glynn F, Alvarez-Iglesias A, Newell J, et al. Effectiveness of a smartphone application to promote physical activity in primary care: the SMART MOVE randomised controlled trial. Br J Gen Pract 2014 Jul;64(624):e384-e391. [doi: 10.3399/bjgp14X680461] [Medline: 24982490]

11. Lee W, Chae YM, Kim S, Ho SH, Choi I. Evaluation of a mobile phone-based diet game for weight control. J Telemed Telecare 2010;16(5):270-275. [doi: 10.1258/jtt.2010.090913] [Medline: 20558620]

12. 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-S137; discussion S138. [doi: 10.1038/sj.ejcn.1602186] [Medline: 16052181]

13. Middelweerd A, Mollee JS, van der Wal CN, Brug J, te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. Int J Behav Nutr Phys Act 2014;11:97 [FREE Full text] [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]

14. Soetens KC, Vandelanotte C, de Vries H, Mummery KW. Using online computer tailoring to promote physical activity: a randomized trial of text, video, and combined intervention delivery modes. J Health Commun 2014 Dec;19(12):1377-1392. [doi: 10.1080/10810730.2014.894597] [Medline: 24749983]

15. Stanczyk N, Bolman C, van Adrichem M, Candel M, Muris J, de Vries H. Comparison of text and video computer-tailored interventions for smoking cessation: randomized controlled trial. J Med Internet Res 2014;16(3):e69 [FREE Full text] [doi: 10.2196/jmir.3016] [Medline: 24589938]

16. Connell D, Goldberg JP, Folta SC. An intervention to increase fruit and vegetable consumption using audio communications: in-store public service announcements and audiotapes. J Health Commun 2001;6(1):31-43. [doi: 10.1080/10810730150501396] [Medline: 11317422]

17. Dijkstra A. Working mechanisms of computer-tailored health education: evidence from smoking cessation. Health Educ Res 2005 Oct;20(5):527-539 [FREE Full text] [doi: 10.1093/her/cyh014] [Medline: 15701665]

18. Hawkins RP, Kreuter M, Resnicow K, Fishbein M, Dijkstra A. Understanding tailoring in communicating about health. Health Educ Res 2008 Jun;23(3):454-466 [FREE Full text] [doi: 10.1093/her/cyn004] [Medline: 18349033]

19. Lustria ML, Noar SM, Cortese J, Van Stee SK, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. J Health Commun 2013;18(9):1039-1069. [doi: 10.1080/10810730.2013.768727] [Medline: 23750972]

20. Oenema A, Brug J, Dijkstra A, de Weerdt I, de Vries H. Efficacy and use of an internet-delivered computer-tailored lifestyle intervention, targeting saturated fat intake, physical activity and smoking cessation: a randomized controlled trial. Ann Behav Med 2008 Apr;35(2):125-135. [doi: 10.1007/s12160-008-9023-1] [Medline: 18363076]

21. Alley S, Jennings C, Persaud N, Plotnikoff RC, Horsley M, Vandelanotte C. Do personally tailored videos in a web-based physical activity intervention lead to higher attention and recall? - an eye-tracking study. Front Public Health 2014;2:13 [FREE Full text] [doi: 10.3389/fpubh.2014.00013] [Medline: 24575398]

22. Lee JA. Effect of Web-based interactive tailored health videos on users' attention, interactivity, overall evaluation, preference and engagement. In: Proc Am Soc Info Sci Tech. 2012 Jan 11 Presented at: American Society for Information Science and Technology; October 9-13, 2011; New Orleans, LA p. 1-3. [doi: 10.1002/meet.2011.14504801317]

23. Chaiken S, Eagly A. Communication modality as a determinant of persuasion: The role of communicator salience. J Pers Soc Psychol 1983;45(2):241-256.

24. Pfau M, Holbert RL, Zubric SJ, Pasha NH, Lin W. Role and influence of communication modality in the process of resistance to persuasion. Media Psychol 2000 Feb;2(1):1-33. [doi: 10.1207/S1532785XMEP0201_1]

25. Corston R, Colman AM. Modality of communication and recall of health-related information. J Health Psychol 1997 Apr;2(2):185-194. [doi: 10.1177/135910539700200215] [Medline: 22013002]

26. World Health Organization. The World Health Report 2002: Reducing Risks, Promoting Healthy Life. Geneva: World Health Organization; 2002.

XSL•FO
RenderX

27. van Rossum CT, Fransen HP, Verkaik-Kloosterman J, Buurma-Rethans EJ, Ocké MC. Dutch National Food Consumption Survey: Diet of Children and Adults Aged 7 to 69 years. (Report No. 35005006/20). Bilthoven: Dutch National Institute for Public Health and the Environment (RIVM); 2011.

28. Netherlands Nutrition Centre. Richtlijnen voedselkeuze (Guidelines for choices of food). 2011. URL: http://www.voedingscentrum.nl/Assets/Uploads/voedingscentrum/Documents/Professionals/Voedselvoorlichting/01_Richtlijnen%20voedselkeuze%20Voedingscentrum.pdf [accessed 2015-08-13] [WebCite Cache ID 6akbkWt3S]

29. Murphy MM, Barraj LM, Spungen JH, Herman DR, Randolph RK. Global assessment of select phytonutrient intakes by level of fruit and vegetable consumption. Br J Nutr 2014 Sep 28;112(6):1004-1018 [FREE Full text] [doi: 10.1017/S0007114514001937] [Medline: 25108700]

30. Kruglanski AW, Thompson EP. Persuasion by a single route: a view from the unimodel. Psychol Inq 2009 Nov 19;10(2):83-109. [doi: 10.1207/S15327965PL100201]

31. Institute of Medicine. Health Literacy: A Prescription to End Confusion. 2004 Apr. URL: http://iom.nationalacademies.org/~/media/Files/Report%20Files/2004/Health-Literacy-A-Prescription-to-End-Confusion/healthliteracyfinal.pdf [accessed 2015-08-13] [WebCite Cache ID 6akcM0hx4]

32. Santo A, Laizner AM, Shohet L. Exploring the value of audiotapes for health literacy: a systematic review. Patient Educ Couns 2005 Sep;58(3):235-243. [doi: 10.1016/j.pec.2004.07.001] [Medline: 16054796]

33. Mõttus R, Johnson W, Murray C, Wolf MS, Starr JM, Deary IJ. Towards understanding the links between health literacy and physical health. Health Psychol 2014 Feb;33(2):164-173. [doi: 10.1037/a0031439] [Medline: 23437854]

34. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]

35. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]

36. Bogers RP, van Assema P, Kester AD, Westerterp KR, Dagnelie PC. Reproducibility, validity, and responsiveness to change of a short questionnaire for measuring fruit and vegetable intake. Am J Epidemiol 2004 May 1;159(9):900-909 [FREE Full text] [Medline: 15105183]

37. TNS NIPO. App bouwen? Begin met Android!. 2011. URL: http://www.tns-nipo.com/nieuws/nieuwsberichten/app-bouwen-begin-met-android!/ [accessed 2015-08-13] [WebCite Cache ID 6akV70ace]

38. Soper D. Statistics Calculator Version 3.0 beta. URL: http://www.danielsoper.com/statcalc3/ [accessed 2015-08-13] [WebCite Cache ID 6akVHLNZ2]

39. Bartholomew LK, Parcel GS, Kok G. Intervention mapping: a process for developing theory- and evidence-based health education programs. Health Educ Behav 1998 Oct;25(5):545-563. [Medline: 9768376]

40. Bartholomew L, Parcel G, Kok G, Gottlieb NH, Fernández ME. Planning Health Promotion Programs: An Intervention Mapping Approach. 3rd edition. San Francisco, CA: Jossey-Bass; 2011.

41. Kok G, Schaalma H, Ruiter RA, van Empelen P, Brug J. Intervention mapping: protocol for applying health psychology theory to prevention programmes. J Health Psychol 2004 Jan;9(1):85-98. [doi: 10.1177/1359105304038379] [Medline: 14683571]

42. Bandura A. Health promotion from the perspective of social cognitive theory. Psychol Health 1998 Jul;13(4):623-649. [doi: 10.1080/08870449808407422]

43. Braverman J. Testimonials versus informational persuasive messages: the moderating effect of delivery mode and personal involvement. Commun Res 2008 Aug 04;35(5):666-694. [doi: 10.1177/0093650208321785]

44. Brug J, Lechner L, de Vries H. Psychosocial determinants of fruit and vegetable consumption. Appetite 1995 Dec;25(3):285-296. [doi: 10.1006/appe.1995.0062] [Medline: 8746967]

45. Springvloet L, Lechner L, Oenema A. Planned development and evaluation protocol of two versions of a web-based computer-tailored nutrition education intervention aimed at adults, including cognitive and environmental feedback. BMC Public Health 2014;14:47 [FREE Full text] [doi: 10.1186/1471-2458-14-47] [Medline: 24438381]

46. Dijkstra A, Schakenraad R, Menninga K, Buunk AP, Siero F. Self-discrepancies and involvement moderate the effects of positive and negative message framing in persuasive communication. Basic Appl Soc Psych 2009 Aug 14;31(3):234-243. [doi: 10.1080/01973530903058441]

47. Lowe MR, Timko CA. What a difference a diet makes: towards an understanding of differences between restrained dieters and restrained nondieters. Eat Behav 2004 Jul;5(3):199-208. [doi: 10.1016/j.eatbeh.2004.01.006] [Medline: 15135332]

48. Adriaanse MA, Vinkers CD, de Ridder DT, Hox JJ, De Wit JB. Do implementation intentions help to eat a healthy diet? A systematic review and meta-analysis of the empirical evidence. Appetite 2011 Feb;56(1):183-193 [FREE Full text] [doi: 10.1016/j.appet.2010.10.012] [Medline: 21056605]

49. Chapman J, Armitage CJ, Norman P. Comparing implementation intention interventions in relation to young adults' intake of fruit and vegetables. Psychol Health 2009 Mar;24(3):317-332. [doi: 10.1080/08870440701864538] [Medline: 20204996]

50. Gollwitzer P, Sheeran P. Implementation intentions and goal achievement. A meta-analysis of effects and processes. Adv Exp Soc Psychol 2006;38:69-119. [doi: 10.1016/S0065-2601(06)38002-1]

51. Boersma P, Weenink D. Praat: doing phonetics by computer (Computer program; Version 5.1.34). 2000. URL: http://www.praat.org [accessed 2015-08-13] [WebCite Cache ID 6akTt7J6A]

XSL•FO

RenderX

52.   Audacity Team. Audacity (Computer program; version 2.0.2). 2012. URL: http://sourceforge.net/projects/audacity/ [accessed 2015-08-13] [WebCite Cache ID 6akUF2gst]

53.   Centraal Bureau voor de Statistiek. Begrippen-Ervaren gezondheid [Definitions-Experienced health]. 2013. URL: http://www.cbs.nl/nl-NL/menu/methoden/begrippen/default.htm?ConceptID=2709 [accessed 2015-08-13] [WebCite Cache ID 6akUWB09Q]

54.   Cohen J, Cohen P, West S, Aiken L. Applied Multiple Regression / Correlation Analyses for the Behavioral Sciences. 3rd edition. Mahwah, NJ: Lawrence Erlbaum; 2003.

55.   Pietersma S, Dijkstra A. Do behavioural health intentions engender health behaviour change? A study on the moderating role of self-affirmation on actual fruit intake versus vegetable intake. Br J Health Psychol 2011 Nov;16(4):815-827. [doi: 10.1111/j.2044-8287.2011.02018.x] [Medline: 21988066]

56.   Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [Medline: 19565683]

57.   Trudeau E, Kristal AR, Li S, Patterson RE. Demographic and psychosocial predictors of fruit and vegetable intakes differ: implications for dietary interventions. J Am Diet Assoc 1998 Dec;98(12):1412-1417. [doi: 10.1016/S0002-8223(98)00319-8] [Medline: 9850109]

58.   Chapman J, Armitage CJ. Do techniques that increase fruit intake also increase vegetable intake? Evidence from a comparison of two implementation intention interventions. Appetite 2012 Feb;58(1):28-33. [doi: 10.1016/j.appet.2011.09.022] [Medline: 22001024]

59.   Brouwer W, Kroeze W, Crutzen R, de NJ, de Vries Nanne K, 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]

60.   Crutzen R, Viechtbauer W, Spigt M, Kotz D. Differential attrition in health behaviour change trials: a systematic review and meta-analysis. Psychol Health 2015 Jan;30(1):122-134. [doi: 10.1080/08870446.2014.953526] [Medline: 25109224]

61.   Lechner L, Brug J, de Vries H. Misconceptions of fruit and vegetable consumption: differences between objective and subjective estimation of intake. J Nutr Educ 1997 Nov;29(6):313-320. [doi: 10.1016/S0022-3182(97)70245-0]

62.   Lathia N, Pejovic V, Rachuri KK, Mascolo C, Musolesi M, Rentfrow PJ. Smartphones for large-scale behavior change interventions. IEEE Pervasive Comput 2013 Jul;12(3):66-73. [doi: 10.1109/MPRV.2013.56]

Abbreviations

BMI: body mass index

XSL•FO
RenderX

Review

Fundamentals for Future Mobile-Health (mHealth): A Systematic Review of Mobile Phone and Web-Based Text Messaging in Mental Health

Sofian Berrouiguet[1], MD; Enrique Baca-García[2], MD, PhD; Sara Brandt[3]; Michel Walter[1], MD, PhD; Philippe Courtet[4], MD, PhD

[1]Brest Medical University Hospital at Bohars, Adult Psychiatry, Brest, France

[2]Department of Psychiatry at Fundación, Jimenez Diaz Hospital, Madrid, Spain

[3]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, USA., New York, NY, United States

[4]Department of Emergency Psychiatry and Post Acute Care, CHRU Montpellier, University of Montpellier, Montpellier, France. FondaMental Foundation, Créteil, France, Montpellier, France

Corresponding Author:
Sofian Berrouiguet, MD
Brest Medical University Hospital at Bohars, Adult Psychiatry
Hôpital de la Cavale Blanche-Bd Tanguy Prigent
Brest, 29200
France
Phone: 33 668204178
Fax: 33 298015218
Email: sofian.berrouiguet@gmail.com

Abstract

Background: Mobile phone text messages (short message service, SMS) are used pervasively as a form of communication. Almost 100% of the population uses text messaging worldwide and this technology is being suggested as a promising tool in psychiatry. Text messages can be sent either from a classic mobile phone or a web-based application. Reviews are needed to better understand how text messaging can be used in mental health care and other fields of medicine.

Objective: The objective of the study was to review the literature regarding the use of mobile phone text messaging in mental health care.

Methods: We conducted a thorough literature review of studies involving text messaging in health care management. Searches included PubMed, PsycINFO, Cochrane, Scopus, Embase and Web of Science databases on May 25, 2015. Studies reporting the use of text messaging as a tool in managing patients with mental health disorders were included. Given the heterogeneity of studies, this review was summarized using a descriptive approach.

Results: From 677 initial citations, 36 studies were included in the review. Text messaging was used in a wide range of mental health situations, notably substance abuse (31%), schizophrenia (22%), and affective disorders (17%). We identified four ways in which text messages were used: reminders (14%), information (17%), supportive messages (42%), and self-monitoring procedures (42%). Applications were sometimes combined.

Conclusions: We report growing interest in text messaging since 2006. Text messages have been proposed as a health care tool in a wide spectrum of psychiatric disorders including substance abuse, schizophrenia, affective disorders, and suicide prevention. Most papers described pilot studies, while some randomized clinical trials (RCTs) were also reported. Overall, a positive attitude toward text messages was reported. RCTs reported improved treatment adherence and symptom surveillance. Other positive points included an increase in appointment attendance and in satisfaction with management and health care services. Insight into message content, preventative strategies, and innovative approaches derived from the mental health field may be applicable in other medical specialties.

KEYWORDS

text messaging; cell phones; mental health; Internet; medical informatics

Introduction

In recent years, the general public and caregivers have increasingly adopted the use of mobile phones. According to the United Nations specialized agency for information and communication technologies, the number of mobile phone subscriptions worldwide had reached almost 7 billion by the end of 2014, corresponding to a penetration rate of 96% worldwide and 90% in developed countries [1]. Mobile health (mHealth) can be defined as the use of mobile computing and communication technologies in health care and public health [2]. mHealth has the potential to incorporate qualities often associated with more conventional health communication methods, such as personalization, tailoring, interactivity, and message repetition at a relatively low cost. Text messaging (short message service, SMS) has proven to be effective, in particular, in psychiatric care. This form of communication allows for the exchange of messages containing 160 characters or fewer between mobile phones [3]. Messages can be sent in a standardized or individualized format and are available on all mobile phones, including low-cost devices. Text message frequency (daily, weekly, etc.), text message interactivity (one-way vs. two-way), personalization (message content based on known characteristics, including patient's condition, history, etc.) and tailoring (message frequency, interactivity and/or content matching each recipient's characteristics) [4]. SMS text messages can also be sent from web-based platforms that allow for pre-scheduling of sending, automation, and better monitoring of reception status.

There is emerging evidence that mobile phones can play an important role in health care delivery, especially in mental health [5]. Combining Internet or mobile phone contact with traditional treatment has shown meaningful results in remote counseling [6] and monitoring support [7]. Literature reviews examining the use of mobile phones in health care have demonstrated the potential of mobile phones to support health education [8], increase access to health care [9], improve prevention and treatment strategies [10], and support public health programs [11]. Text messaging has also been used to provide appointment reminders [11], improve patient adherence with treatment [12], monitor chronic conditions [13], and provide psychological support [14]. Additionally, SMS messages are used in the prevention of communicable diseases [15] and in preventive health promotion programs [16]. Text messaging has also improved service provision to population subgroups who do not typically use health services [13]. Text messages provide the opportunity to remotely access caregivers for advice [17], and mHealth can also extend prevention strategies for caregivers; another on-going challenge of 21st century medicine [18]. Knowledge proceeding from the mental health field could also be easily transferred to other specialties, considering the transversal contribution of cellular technology in providing innovative monitoring and prevention strategies.

When taking into account the rapid expansion of mHealth applications, combined with current mental health challenges, reviews of existing applications and evidence of successful messaging-supported health care strategies are needed. Reviews of text messaging have been proposed in many fields of medicine including diabetes, weight loss, and smoking cessation [4]. To better understand and validate the use of these specific mHealth applications, we conducted a review of the use of mobile phone text messaging to promote mental health.

Methods

Objective

The objective of this study was to provide a thorough review of the applications of text messaging in mental health care. Two review authors independently assessed all studies retrieved against the inclusion criteria (SB and EBG). Disagreements were resolved by a third review author (PC). During the literature review process, relevant studies were categorized in a two-step approach. We first performed the review of the titles and abstracts of all publications that were identified as relevant according to the inclusion criteria. Abstracts were then categorized by the type of methodology used, health condition, applications, and purposes. The full text of all publications that were not excluded during the title and abstract review stage were checked. Publications that met all inclusion criteria comprised the final sample.

Search Strategy

PubMed, PsycINFO, Cochrane, Scopus, Embase and Web of Science databases were extensively searched in May 2015. The list of keywords was created around the two domains of *mental health* and *text messaging*. A search command was constructed using "AND" and the disjunction "OR" as logical operators in Medical Subject Headings (MeSH) terms, titles, and abstracts (see Figure 1). We did not include keywords related to smoking cessation as it had already been reviewed in another article [19]. Substance abuse disorders were included in our review despite a recent study that did not take into account recent articles and pilot studies [20]. We explored titles, abstracts, and MeSH terms, and reference lists of selected studies were also checked for other potentially relevant studies.

Primary endpoints of interest were health outcomes as a result of text messaging. Technical aspects were sometimes reported, and we also considered the patients' and caregivers' evaluation of messaging approaches, subjective perceptions of effectiveness, institutional burden, and cost when possible. The studies included were heterogeneous in terms of design features, conditions addressed, characteristics of messaging procedure, and outcome measures. Thus, the findings are presented descriptively.

Figure 1. Flowchart of review process.
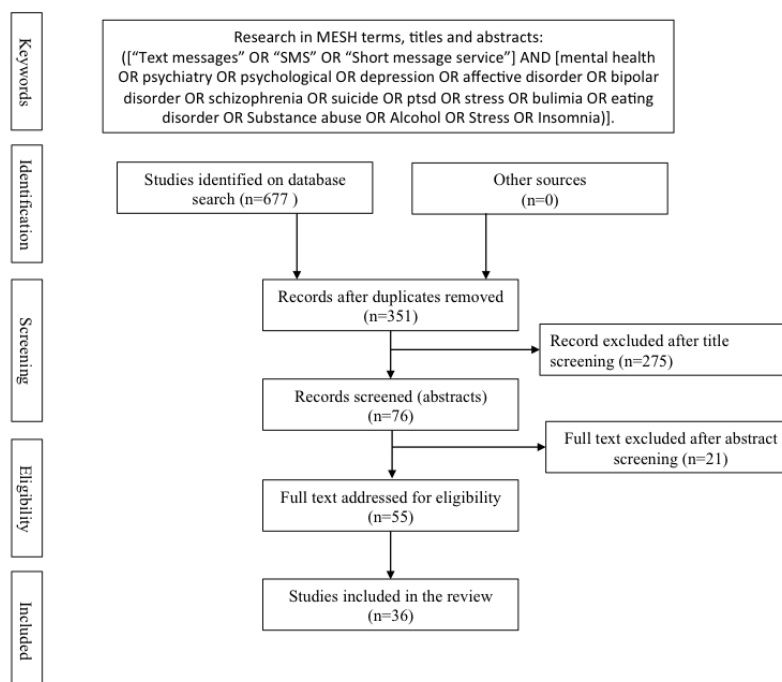
Selection Criteria

We included randomized controlled trials (RCTs) and non-randomized studies. Study protocols were included when referring to RCTs. We included studies in which SMS text messaging was used to promote mental health, including any type of preventive or monitoring strategy. Text messages could be delivered to a patient by the caregiver or vice versa. We did not include studies assessing the patient's general feeling about text messages (ie, surveys) for a specific medical application. We excluded studies that used mobile phone multimedia messaging service or android/OSX apps as tools for prompting.

Results

The steps in the literature search and review process are summarized in Figure 1. The initial search retrieved 677 articles. After checking for duplicates and screening abstracts and full texts, 36 articles met the inclusion criteria. Notably, most studies (20/36, 56%) assessed the feasibility and acceptability of sending text messages to patients with a psychiatric condition or substance abuse disorder, while 44% (16/36) of the studies proposed quantitative methodologies. Ten of the 36 studies were RCTs (28%).

We observed increased interest in text messaging, from 4 articles cited in 2010 to 14 articles cited in 2014. The largest upsurge occurred between 2013 and 2014 (see Figure 2). The majority of publications regarding text messaging were in medical journals (29/36, 81%), with the remaining 19% (7/36) being published in journals focused on medicine and telecommunication. We did not retrieve any articles published in journals focused only on technology. The geographical areas involved in this type of research were mainly Europe (16/36, 44%) and the United States (17/36, 47%).

Health Conditions

Studies addressed a wide range of psychiatric conditions. Figure 3 shows the mental health conditions in which the use of text messaging was studied. Substance abuse conditions were most often studied (11/36, 30%), followed by schizophrenia (8/36, 22%), and affective disorders, including bipolar disorder and depression (6/36, 17%). We found only eight studies on suicidal behavior, eating disorders, and post-traumatic stress disorder (PTSD). Text messaging was used in the management of chronic conditions (27/36, 75%), reactive conditions (7/36, 19%), and preventative strategies for healthy or at-risk individuals (2/36, 6%).

Figure 2. Distribution of articles over time.
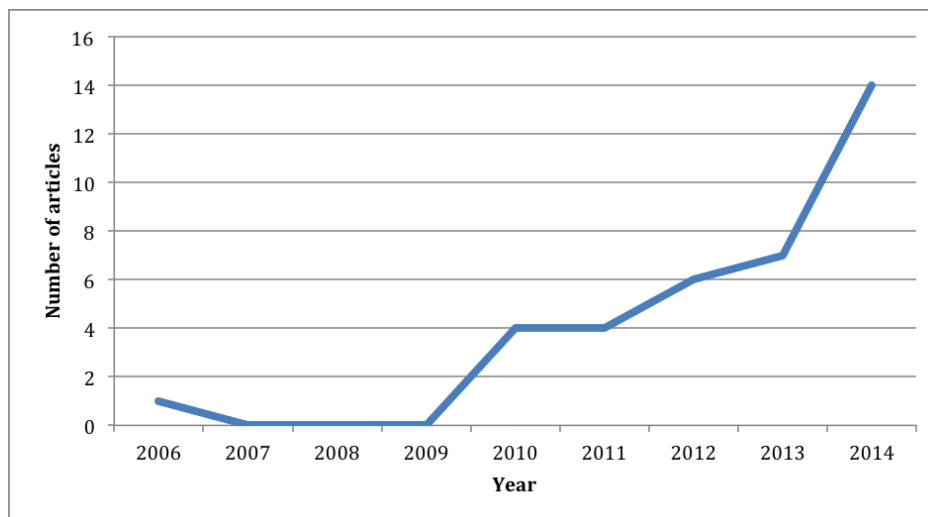
Figure 3. Mental health conditions addressed by text messaging.
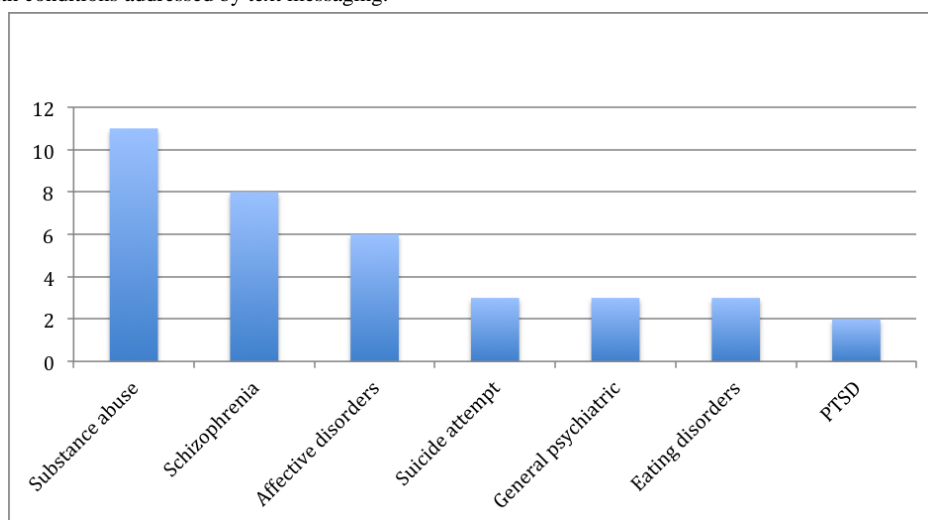
Methodology

Concerning the methodology of the studies, we identified pilot studies (20/36, 56%), followed by RCTs (10/36, 28%). Three of the RCT articles described only the study protocols with no results. Six studies used non-randomized, comparative methods. Before 2013, the samples used to study text messaging were mostly small (ie, fewer than 80 patients). Feasibility and acceptability were assessed in small sample sizes (ie, 15 to 50 patients) **.**The sample size increased significantly in 2013, with four RCTs assessing samples of more than 400 patients.

Application

We identified four uses of text messages: reminders (5/36, 14%), information provision (6/36, 17%), supportive messages (15/36, 42%), and self-monitoring procedures (15/36, 42%). Applications were sometimes combined. Reminders aimed to improve either appointment attendance or treatment adherence. Information aimed to provide important notifications concerning available health care services or health recommendations. Supportive messages were either tailored to the patient or standardized. For self-monitoring procedures, patients had to send messages on their situation in time increments ranging from once weekly to twice daily.

Purposes

Most studies assessed the feasibility and acceptability of text messaging for patients. Outcomes were assessed by questionnaires on satisfaction or recorded response rates. We found only one study out of 36 that concluded that messaging was not well accepted, which involved a sample of patients with eating disorders [21]. Apart from this article, the other 35 studies reported positive perception of text messaging on the primary outcomes, which were acceptability, attendance at appointments, treatment adherence, and improvement of health.

Substance Abuse

SMS text messaging was frequently used to assess patients with issues related to substance abuse, as summarized in Table 1. A three-arm randomized trial compared self-reported alcohol use three months after emergency department visits, during which 765 young adults reported hazardous drinking [22]. Weekly text message drinking assessments were sent. The number of self-reported binge drinking days decreased from baseline to three months in the real-time feedback group compared to text

message drinking assessments without feedback or a control condition. Moore et al [23] assessed the feasibility of using text messages both to survey and moderate alcohol use, and were able to determine periods of greater alcohol use (ie, weekends and celebratory events). Irvine et al [24] also used two-way text messaging as a preventative method to assess and support disadvantaged men at risk of substance abuse. Interview data indicated that text messaging was acceptable to participants, and preferred over email and web-based methods. Text messaging also proved to be accepted and effective for young people (ages 12 to 24) transitioning out of community-based substance abuse treatment programs over a 90-day period [25]. In addition to alcohol abuse, other researchers performed ecological momentary assessments (EMAs) via text messaging on daily methamphetamine use, craving levels, and the perceived usefulness of messages [26]. The odds of messages being rated as *very* or *extremely* useful were 6.6 times higher (95% CI 2.2-19.4) in the active periods than the placebo periods, signifying good acceptability for text messages.

Kuntshe and Robert [27] showed that text message reports were as accurate as reports by mail. Other teams explored perspectives concerning the appeal, acceptability, and content of text messaging in a population at risk of alcohol abuse after discharge from the emergency department [22,28], and in adolescents admitted to primary care clinics [29].

Studies also reported the use of supportive messages in substance abuse. Bedtsten et al [30] compared two methods of delivery by randomizing young students to receive automated alcohol-intervention messages, either by text messaging or by email. No difference was reported regarding satisfaction with the length and frequency of the messages, regardless of the method of delivery. Similar results were reported by Gonzales et al in young people [31]. In a longitudinal study, Haug et al [32] showed decreases in the percentages of persons with risky single-occasion drinking from baseline to follow-up assessment (75.5% vs. 67.6%, *P*<.001) in 364 students. A pilot study also examined the feasibility of a 12-week mobile-based aftercare program for youth (ages 12 to 24) transitioning out of community-based substance abuse treatment programs over a 90-day period [25]. A study protocol described the use of text messaging in heavy drinkers [33] in order to provide initial information about the feasibility and efficacy of mHealth interventions for improving treatment adherence in alcohol use disorders.

Schizophrenia

Seven of the 36 studies focused on a diverse range of text message interventions for patients suffering from schizophrenia or related disorders, as summarized in Table 2. Two studies proposed text message reminders as a means to improve treatment adherence in patients with schizophrenia [34,35]. Montes et al [33] performed a prospective, randomized, open-label, controlled, six-month study in 56 outpatient psychiatric centers. Participants assigned to the intervention group received daily text messages on their mobile phones for three months, reminding them to take their medication consistently. The text message said, "*Please remember to take your medication*". A significant improvement in adherence was

observed among patients receiving text messages compared to the control group. Pijnenbrog et al [35] also found that patients significantly improved in keeping appointments with mental health workers when sent text message reminders, and carrying out leisure activities also increased with text message reminders.

Two studies examined text message support in outpatient management of patients with psychosis [36,37]. In the study by Ben-Zeev et al [37], the objective was to assess response rates to treatment-adherence messages. Results showed high rates of prompt responses and satisfaction rates (up to 90%) toward text message interventions. Valimaki et al [36] presented a study protocol in which the patient chose the form, content, timing, and frequency of the messages.

Symptom assessment via text messaging was proposed to patients with schizophrenia in three studies [38-40]. Ganholm et al [40] and Depp et al [38] presented two versions of the same project: Mobile Assessment and Treatment for Schizophrenia. Through this method, these studies were able to assess self-reported treatment adherence, the number of social interactions, and severity of auditory hallucinations over a 12-week period. Overall, both studies showed that text messaging for symptom monitoring is a feasible and effective form of intervention to alleviate symptoms in patients with schizophrenia. Ainsworth et al [39] compared two methods of assessing symptoms via mobile phone, and demonstrated that it was feasible using either text messages or native smartphone applications.

Affective Disorders

Text message interventions were proposed to patients with affective disorders, including depression and mania [41-46], as summarized in Table 2. Supportive messages are used to convey positive feelings from the caregiver to patients with affective disorders. Agyapong et al [47] sent supportive twice-daily messages to a sample of 54 patients over a 3-month period. During this period, patients in the intervention group had significantly lower scores on the Beck Depression Inventory-II than control groups (8.5, standard deviation [SD] 8.0 vs. 16.7, SD 10.3, *P*=.003) [44]. Aguilera [46] also used supportive text messaging in adjunct to cognitive-behavioral therapy for depression in two English-and Spanish-native speaking populations. The findings proved the interventions were feasible and acceptable to patients and, for Spanish-speaking patients, were also related to increased feelings of being cared for.

Symptom assessment was proposed by Bopp et al [42] to assess the course of bipolar disorder symptoms through a weekly assessment with the Altman Self-Rating Mania Scale and the Quick Inventory of Depressive Symptoms-Self Report. Adherence with the procedure was high (75%) and participants more frequently reported depressive symptoms (47.7%) than manic symptoms (7%).

Eating Disorders, Suicidal Behavior, and Post-Traumatic Stress Disorder

Text messaging has also proven to be acceptable and effective in a number of other mental disorders, such as eating disorders, suicidal behavior, and PTSD. Studies regarding text message interventions in patients with these conditions are summarized

XSL•FO

RenderX

in Table 3. Three studies aimed to determine whether text message interventions might be of use in the management of eating disorders. Robinson et al [21] showed a low response rate to the self-monitoring procedure, suggesting that the intervention was only moderately well accepted by participants. However, Shapiro et al [48] showed that 87% of participants adhered to self-monitoring and reported good acceptability. Bauer et al [49] added tailored feedback to self-monitoring. In an RCT, this study reported that text messaging could support significant improvement in remission rates after discharge from inpatient treatment for bulimia nervosa. Two studies explored the feasibility of post-acute crisis text messaging outreach for suicide prevention. Weekly text messages were sent to patients

after a suicide attempt. Efficacy on further suicide attempts is currently being assessed in an RCT. In both studies, messages were identical for all patients. Berrouiguet et al [50] used a web-based application in order to provide personalized and tailored messages, and to manage a larger number of patients with the end goal of reducing the rate of suicide reattempts.

Studies relating to PTSD management showed that text messaging is applicable for support of at-risk patients [51] and for monitoring acute symptoms [52]. Other messaging interventions were proposed to patients during routine follow-up, without targeting any specific disorder. These applications were essentially appointment reminders for outpatients [53,54] and preventive information for disadvantaged patients [55].

Table 1. Summary of text message interventions in patients with substance abuse.

Authors (year)	Country	Population (sample)	Text messages	Principal outcome	Method, duration	Result
Suffoletto (2011)	USA	Young adults from urban emergency departments (n=45)	Text messages self-monitoring	Feasibility of heavy drinking days and drinks per drinking days, assessment by text message.	Randomized comparative study	Feasible
Stoner et al (2012)	USA	Treatment-seeking heavy drinkers (expected n=105)	Medication reminders and assessment	Effectiveness	Randomized trial	To be published
Haug et al (2013)	Switzerland	Vocational school students (n=477)	Self-monitoring	To evaluate appropriateness	Longitudinal pre-post study	Study found reduced percentage of persons with risky single-occasion drinking from baseline (75.5%, 210/278) to follow-up assessment (67.6%, 188/278, $P<.001$)
Keoleian et al (2013)	USA	Methamphetamine users (n=5)	Self-monitoring	Feasibility	Randomized crossover pretest pilot study	79% of scheduled assessment were collected.
Mason et al (2013)	USA	College students with alcohol problems (n=18)	Self-monitoring and supportive messages	Feasibility and effectiveness	Randomized trial	Text messages for alcohol abuse prevention are feasible.
Rios-Bedya et al (2013)	USA	Adolescents recruited in primary care clinics (n=29)	Ecological momentary assessment	Feasibility	Pilot study	High participation rate
Bendsten et al (2014)	Sweden	University students (n=454)	Self-monitoring and supportive messages	Satisfaction regarding text messages	Randomized trial	No difference was seen regarding satisfaction with length and frequency of messages, regardless of method of delivery.
Lucht et al (2014)	Germany	Inpatient after alcohol detoxification (n=80)	Information about telephone support, twice a week.	Controlled prospective open pilot study.	Pilot study	Feasibility and acceptability were good. Adherence was satisfactory with 57.14% of participants replying to at least 50% of prompts.
Moore et al (2014)	UK	Alcohol consumers recruited in university (n=80)	Self-monitoring	Acceptability	Randomized controlled trial	Acceptable and preferred to email conducted assessment
Rachel Gonzales et al (2014)	USA	Young participants transitioning out of substance abuse program (n=80)	Self-monitoring, supportive messages	Feasibility	Random	A significant effect of condition on primary drug use relapse outcomes over time was observed as measured by urine analysis.
Suffoletto et al (2014)	USA	Young adults discharged from emergency department (n=765)	Self-monitoring	Satisfaction towards text message or email contact	Randomized trial	Decreased number of binge drinking in web intervention group only

Table 2. Summary of studies using text messages in patients with schizophrenia and affective disorders.

Authors (year)	Country	Population (sample)	Text messages	Principal outcome	Method, duration	Result
Pijenborg et al (2010)	Netherlands	Patients suffering from schizophrenia with severe cognitive impairment (n=62)	Text message reminders	Improvement in functioning in daily life	Non-randomized controlled trial, 7 weeks	The overall percentage of goals achieved increased with prompting (eg, appointments). Patients enjoyed receiving the message.
Depp et al (2010)	USA	Patients with severe mental illness (n=8) in pilot study program	Text messages for EMA	Feasibility evaluation	Pilot study	Monitoring symptoms of patients suffering from severe mental illness using text messages is feasible.
Granholm et al (2012)	USA	Patient suffering from schizophrenia (n=55)	Self-monitoring	Responding rate	Pilot study, 12 weeks	Text messaging interventions are feasible and effective in patients with schizophrenia.
Maritta Välimäki et al (2012)	Finland	Patient with psychosis (protocol)	Text message support	To evaluate the impact of text messages to encourage treatment adherence and follow-up	Randomized trial, 12 months	To be published
Montes et al (2012)	Spain	Patients suffering from schizophrenia (n=254)	Daily text message reminders	Impact of text messages on adherence with antipsychotic treatment	Multicenter, randomized, open-label, controlled study, 3 months	Significant improvement in adherence
Ainsworth et al (2013)	UK	Patients with non-affective psychosis (n=24)	Symptom assessment via text messages or native smartphone application (EMA)	Compare text message based assessment strategies to native mobile texting application in terms of satisfaction	Randomized, repeated measure, crossover design, 3 weeks	A greater proportion of data points were completed with the native smartphone application.
Bebee et al (2014)	USA	Outpatients followed for schizophrenia (n=30)	Self-monitoring	To evaluate the impact of text messages on treatment adherence	Comparative study with random assignation in intervention groups, 3 months	Non-significant effect on treatment adherence
Ben-Zeev et al (2014)	USA	Patients with psychotic disorder and substance abuse (n=70)	Text message support	Feasibility and acceptability	Pilot qualitative study, 12 weeks	90% of patients found the intervention useful

Table 3. Summary of text messaging in outpatients with PTSD, suicide attempters, and patients with anorexia and/or bulimia.

Authors (year)	Country	Population (sample)	Text intervention	Principal outcome	Method, duration	Result
Chandra et al (2014).	India	Girls in the age range of 16-18 years from urban slums (n=40)	Information messages about health promotion	Feasibility and acceptability	Pilot qualitative study, 1 month	Mobile text messages are a feasible and culturally acceptable method for mental health promotion.
Kunigiri et al (2014).	UK	Psychiatry outpatients (n=2556)	Text message reminders	Attendance at follow-up appointment	Three-arm comparative study, text messages sent 14 days and 2 days prior to appointment	Significant increase in the attendance in text message reminder group compared to telephone
Branson CE et al (2013).	USA	Psychiatry outpatient (n=48)	Text message reminders	Text message reception rate	Pilot study examinating technical feasibility	Patients received 88% of scheduled text messages. High patient satisfaction reported.
Price et al (2014).	USA	Patient suffering from PTSD (n=29)	Self-monitoring	Responding rate	Pilot study, 3 months	Text message described as a viable method to monitor PTSD.
Furber et al (2014)	Australia	Patient presenting to the Emergency Department for emotional crisis (n=68)	Supportive messages	Text message intervention acceptance rate	Non-randomized comparative prospective study (compared to historical control group, 6 months	66% of patients accepted the intervention. No significant differences in clinical outcomes between groups.
Chen et al (2011)	China	Patients discharged after suicide attempt (n=15)	Supportive messages	Feasibility and acceptability of messages to attempters after discharge	Pilot study, 1 month	Seen as feasible and acceptable to suicide attempters. Showed desire to keep receiving message.
Berrouiguet et al (2014)	France	Patients discharged after suicide attempt (n=15)	Supportive and information messages	Feasibility and acceptability	Pilot study, 12 months	Suicide attempters accepted text messages.
Berrouiguet et al (2014)	France	Patients discharged after suicide attempt (n=520)	Supportive and information messages	Suicide reattempts	Study protocol of randomized controlled study	To be published
Robinson et al (2006)	UK	Patients with bulimia nervosa (n=21)	Self-monitoring	Acceptability and feasibility	Pilot study, six months	Low participation rate and high attrition rate
Shapiro JR et al (2010)	USA	Patients with bulimia nervosa (n=31)	Self-monitoring	Participation rate	Pilot study, six months	87% of participants adhered to self-monitoring.
Lucht et al (2014)	Germany	Patient with bulimia nervosa (n=165)	Self-monitoring	Impact of text messaging on remission rate after 8 months	Randomized controlled trial, 16 weeks	Text messaging improved remission rate in intervention group (51%) compared to control group (36.1%).

Technological Aspects

Regarding technological aspects, text messages were sent from a web-based application or a mobile phone (see Multimedia Appendix 1). In most studies, patients were encouraged to send text messages to caregivers for monitoring purposes (25/36, 69%). In only two studies (2/36, 6%), the patients were the only individuals to send text messages. Eleven studies proposed a one-way caregiver to patient text message exchange. In some studies, mobile devices were provided to patients for the study period [40,54]. One study also used an original smartphone application to manage text messaging [39].

Discussion

Text messaging is an effective means to assess the impact of mHealth interventions in unbiased samples and in a widespread population. Only 32% of the population worldwide own a smartphone [1]. In comparison, ubiquitous access to text messaging in the general population has recently been reached, making it a much more feasible method for managing psychiatric disorders in the broadest population.

We extensively screened all published papers dealing with the use of text messaging in the field of mental health care. Despite

recent rapid growth, this innovative approach is still at an early stage. It remains difficult to interpret research findings because of heterogeneous study designs, populations, and medical conditions. Providing a review that could effectively translate findings into best-practice methods remains impossible, as has been described in other reviews concerning mHealth [4]. We aimed to descriptively report the research that has been conducted in the past decade concerning text messaging in mental health care in terms of medical conditions, characteristics of the interventions, and outcomes.

Results from this literature review demonstrate that text messaging as a management tool for mental health has been proposed in many mental health conditions with promising results. Innovative strategies have been suggested, but there is still a lack of evidence regarding their efficacy due to the paucity of consistent RCTs. We found a growing number of initiatives toward incorporation of text messaging in existing mental health care strategies, most of which were pilot studies. The increase in the number of articles over the past decade also indicates growing interest in the topic in peer-reviewed scientific literature. In particular, the number of articles nearly doubled from 2013 to 2014. These results are consistent with the increasing use of text messaging in all fields of health care management. Text messaging is highly adaptable to any health care strategy, given that it does not interfere with pre-existing care procedures. The findings from these text messaging studies may contribute to innovation in other fields of medicine as well.

mHealth and the Emergence and Expansion of Ecological Momentary Assessment

Although there are still a number of ways in which text messaging can be improved, it is clear that there are many overarching benefits to the use of mHealth techniques in the clinical field. With increasing availability of technology such as mobile phones and smartphones, not only are new forms of clinical intervention possible, but EMA has also arisen as a powerful research tool, especially in younger populations. Self-assessment and EMA have produced promising results when used in patients with mental disorders. EMA extends the concept of self-monitoring to emphasize real-world, real-time data capture. The ability to reach out to patients on a regular basis via text messaging allows for more effective use of EMA since patients can be assessed in their natural environments, rather than in a hospital or counseling setting. This advantage allows for a more accurate and comprehensive reading of patients' physical and mental status, thus enabling physicians to better treat patients.

However, little is known about the validity provided by the assessment. In our review, researchers often used custom assessment procedures. An interesting approach, proposed by Altman et al (2003), provides patients with a self-assessment method relying on a validated self-assessment scale. This consideration may be of importance given that other medical specialties also intend to implement EMA (eg, in diabetes and asthma management) and may face the same validity issues [5]. Migration of such methods to an automated environment requires prior validation [7].

Insight Into the Type of Text Messaging Used

Overall, text messaging was used for reminders, information provision, supportive messages, and self-monitoring. Messages were never sent as a substitute for consulting or treatment. In our review, text messaging was proposed to patients suffering from either acute or chronic conditions. Text messaging invariably provided an extension of traditional care strategies, and this extension may occur after discharge [22], in between counseling sessions, or for delivery of a preventive message to an at-risk population [32]. Even when text messaging was used to perform a single assessment, it had a positive effect on clinical outcomes [49]. This finding is an important point that highlights a common aim in mHealth strategies: the strengthening of continuity in existing care. Text messaging in mental health seems to embrace a broad panel of messaging possibilities, breaking away from an approach that would rely solely on reminders [56]. These findings may encourage initiatives in other fields of medicine to assess the impact of text messaging and all of its available features.

Managing Patient Refusal

We found that text messaging may be readily accepted, even in populations that might normally be reluctant or opposed to treatment. The attitude of patients toward mental health services is crucial to seeking help, determining pathways of care, and subsequent therapeutic commitment and adherence. In fact, mental health care services often have to deal with patient refusal. Text messaging, however, was well accepted by patients suffering from eating disorders (a population that often refuses traditional treatment), as well as suicide attempters who accepted text messaging after refusing hospitalization [57]. Similar challenges can be overcome through the use of messaging in immunization campaigns [58], insulin therapy [59], and cancer management [60]. Text messaging is able to renew the care-giving process and avoids disruptions in the patient-physician relationship. Text messaging may also be an opportunity for patients to retain access to important health information without the potential stigma associated with clinic visits.

Reduction of Social Isolation and Increased Patient Interactivity

Text messaging is also particularly useful to reduce social isolation. As with other chronic conditions, mental illness is typically associated with increased social isolation, and text messaging may be an important tool to combat this issue [61]. Social isolation can be defined as disengagement from social ties, institutional connections, and even access to care services. Traditionally, interventions that provide social support for patients with disabling conditions rely on caregivers [62] or outpatient nursing programs [63]. The mechanism for maintaining patient participation may be as simple as a weekly message asking, "*How are you?*" [50] or, "*Thank you*" messages [47]. These spontaneous forms of contact allow patients to self-report their problems to caregivers using text messages, or to seek telephone or face-to-face contact. Two-way text messaging may also allow patients to become more involved in treatment, as shown in patients with asthma or diabetes [3,4]. Overall, it is clear that text messaging can provide effective

two-way communication and promising support to enhance patient involvement and interactivity, and ultimately to reduce social isolation.

Limits and Recommendations

In our study, we propose a review of mHealth interventions based on text messaging. Despite the many benefits of SMS text messaging, there are still some limitations to the use of text messages in the clinical realm, which need to be addressed. Since the inception of text messaging, other innovative mHealth strategies have been developed using recent technological advances [64]. Due to the simplicity of its content, text messaging cannot be used as a remote counseling tool, unlike other telemedicine devices [6]. However, even with a few words, a simple message can have an important impact. The content of messages is of particular importance. Some characteristics such as personalization, caring sentiments, and polite text are associated with more successful preventative messages [65]. Most texting applications in mental health share a common characteristic: they tend to enhance connectedness between patient and caregivers. This concept was introduced before the mHealth era. In a pioneer study conducted in the 1970s, Motto proposed to stay in touch with patients discharged after a suicide attempt through regular surface mail for five years [66]. This effect has been described as testimony of the concern that caregivers have for the patient's situation. Unfortunately, we did not retrieve any studies that addressed the question of message content. Furthermore, automated sending (which was often used in the included studies) was not conducive to basic precautions regarding message content. More concern and better knowledge is required in this area.

Another limitation present throughout the studies was that baseline behaviors with regard to the use of cellular technology varied widely between population subgroups [1,67]. We found only one study that assessed the use of mobile phones before enrolment in the study [37]. Since important variations have been described depending on population subgroups, better description of the population studied could help identify which patients would most likely respond favorably to text messaging. In the general population, the desire to use mobile phones also varies depending on age [68] and gender [69]. A better description of sociodemographic patterns may also be of crucial importance in selecting the right intervention for the right subgroup. Other factors such as level of education seemed to be associated with a better response to preventive messaging [70]. Such idiosyncratic factors may help predict the effects of text messaging. Insight into these factors likely requires better assessment of baseline behaviors.

Several important ethical considerations related to the use of text messaging in mental disorders must also be considered. The matter of participant burden is one such issue. Text messaging entails a non-negligible time commitment on behalf of the participants, and some text messaging programs rely on

daily, or sometimes more frequent, prompts. Studies usually last for several days [52] or months, often after discharge or between counseling sessions. Furthermore, recording participants' daily experiences in a continuous manner is an integral part of EMA. This approach may be significantly more invasive than asking a participant to complete a retrospective questionnaire or answering a question at a traditional interview. The risk of intrusiveness into daily life is real, and yet this issue was not assessed in the articles we reviewed or in other reviews in the field [71]. Receiving a text message may also inconvenience participants who are expected to complete an EMA or read information at a moment's notice (ie, requests may occur at inopportune times). Caregivers should ascertain that such a burden would not be detrimental to participants' well-being, particularly when studying individuals who have recently remitted or who are in-episode.

The number of messages can also vary based on the patients' condition and target behaviors. Some patients may need to receive weekly messages (eg, in suicide management [41]), whereas other patients may need to receive daily messages to remind them to take medication [34] or for monitoring purposes [38]. Pop-Eleches et al [72] tested both daily and weekly messages and found that weekly messages improved anti-retroviral therapy adherence whereas daily messages did not. The frequency of text messages has been described as an important factor for success, as well as failure. In reminding patients to take their medication, weekly messages were more likely to achieve adherence above 95%. Daily messages were likely to be intrusive and cause user fatigue, thereby rendering them ineffective. Further research is necessary to understand the role of timing in the efficacy of text messaging, taking into account the potential intrusiveness of such interventions.

Conclusion

Overall, it is clear that text messaging has numerous benefits, from extending EMA and increasing patient interactivity to improving mental health in patients with chronic conditions, and encouraging treatment in normally resistant populations. Text messages and mHealth interventions are at a crucial stage in their development, as they present a promising opportunity for innovation in medicine, especially in terms of connectedness between patients and care services. At the same time, the risk of intrusiveness linked to the entry of care services into a patient's personal space is high. This risk should be more carefully assessed. Nevertheless, text messages allow for inexpensive and instantaneous communication between patients and clinicians, and remains the easiest way to access mHealth applications. All results could be transferred to prompting applications using smartphone technologies. Additionally, early studies suggest that text messaging may be helpful for treatment adherence. Text messaging could also be useful for other aspects of patient self-management, by enhancing social support, encouraging patients to become more proactive in health care, and providing information to enhance health and well-being.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Technical features of text-messaging procedure.

[PDF File (Adobe PDF File), 31KB - jmir_v18i6e135_app1.pdf]

References

1. International Telecommunication Union. ICT facts and figures-the world in 2015. 2015. URL: http://www.itu.int/en/ITU-D/Statistics/Pages/facts/default.aspx.(Archived [accessed 2015-04-29] [WebCite Cache ID 6atZAxj1W]

2. Free C, Phillips G, Felix L, Galli L, Patel V, Edwards P. The effectiveness of m-health technologies for improving health and health services: a systematic review protocol. BMC Res Notes 2010;3:250 [FREE Full text] [doi: 10.1186/1756-0500-3-250] [Medline: 20925916]

3. Lim MS, Wright C, Hellard ME. The medium and the message: fitting sound health promotion methodology into 160 characters. JMIR Mhealth Uhealth 2014;2(4):e40 [FREE Full text] [doi: 10.2196/mhealth.3888] [Medline: 25367387]

4. Hall AK, Cole-Lewis H, Bernhardt JM. Mobile text messaging for health: a systematic review of reviews. Annu Rev Public Health 2015 Mar 18;36:393-415 [FREE Full text] [doi: 10.1146/annurev-publhealth-031914-122855] [Medline: 25785892]

5. Fiordelli M, Diviani N, Schulz PJ. Mapping mHealth research: a decade of evolution. J Med Internet Res 2013;15(5):e95 [FREE Full text] [doi: 10.2196/jmir.2430] [Medline: 23697600]

6. Hilty DM, Ferrer DC, Parish MB, Johnston B, Callahan EJ, Yellowlees PM. The effectiveness of telemental health: a 2013 review. Telemed J E Health 2013 Jun;19(6):444-454 [FREE Full text] [doi: 10.1089/tmj.2013.0075] [Medline: 23697504]

7. Mundt JC, Greist JH, Gelenberg AJ, Katzelnick DJ, Jefferson JW, Modell JG. Feasibility and validation of a computer-automated Columbia-Suicide Severity Rating Scale using interactive voice response technology. J Psychiatr Res 2010 Dec;44(16):1224-1228. [doi: 10.1016/j.jpsychires.2010.04.025] [Medline: 20553851]

8. Sahu M, Grover A, Joshi A. Role of mobile phone technology in health education in Asian and African countries: a systematic review. Int J Electron Healthc 2014;7(4):269-286. [doi: 10.1504/IJEH.2014.064327] [Medline: 25161104]

9. Townsend A, Leese J, Adam P, McDonald M, Li LC, Kerr S, et al. eHealth, participatory medicine, and ethical care: a focus group study of patients' and health care providers' use of health-related Internet information. J Med Internet Res 2015;17(6):e155 [FREE Full text] [doi: 10.2196/jmir.3792] [Medline: 26099267]

10. Zurovac D, Sudoi RK, Akhwale WS, Ndiritu M, Hamer DH, Rowe AK, et al. The effect of mobile phone text-message reminders on Kenyan health workers' adherence to malaria treatment guidelines: a cluster randomised trial. Lancet 2011 Aug 27;378(9793):795-803 [FREE Full text] [doi: 10.1016/S0140-6736(11)60783-6] [Medline: 21820166]

11. Kannisto KA, Koivunen MH, Välimäki MA. Use of mobile phone text message reminders in health care services: a narrative literature review. J Med Internet Res 2014;16(10):e222 [FREE Full text] [doi: 10.2196/jmir.3442] [Medline: 25326646]

12. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]

13. Levy N, Moynihan V, Nilo A, Singer K, Bernik LS, Etiebet MA, et al. The Mobile Insulin Titration Intervention (MITI) for insulin adjustment in an urban, low-income population: randomized controlled trial. J Med Internet Res 2015;17(7):e180 [FREE Full text] [doi: 10.2196/jmir.4716] [Medline: 26187303]

14. Agyapong VI, Mrklas K, Suen VY, Rose MS, Jahn M, Gladue I, et al. Supportive text messages to reduce mood symptoms and problem drinking in patients with primary depression or alcohol use disorder: protocol for an implementation research study. JMIR Res Protoc 2015;4(2):e55 [FREE Full text] [doi: 10.2196/resprot.4371] [Medline: 25979786]

15. Sheoran B, Braun RA, Gaarde JP, Levine DK. The hookup: collaborative evaluation of a youth sexual health program using text messaging technology. JMIR Mhealth Uhealth 2014;2(4):e51 [FREE Full text] [doi: 10.2196/mhealth.3583] [Medline: 25367444]

16. Abroms LC, Johnson PR, Heminger CL, Van Alstyne JM, Leavitt LE, Schindler-Ruwisch JM, et al. Quit4baby: results from a pilot test of a mobile smoking cessation program for pregnant women. JMIR Mhealth Uhealth 2015;3(1):e10 [FREE Full text] [doi: 10.2196/mhealth.3846] [Medline: 25650765]

17. Li Y, Wang W, van Velthoven MH, Chen L, Car J, Rudan I, et al. Text messaging data collection for monitoring an infant feeding intervention program in rural China: feasibility study. J Med Internet Res 2013;15(12):e269 [FREE Full text] [doi: 10.2196/jmir.2906] [Medline: 24305514]

18. Adelman RD, Tmanova LL, Delgado D, Dion S, Lachs MS. Caregiver burden: a clinical review. JAMA 2014 Mar 12;311(10):1052-1060. [doi: 10.1001/jama.2014.304] [Medline: 24618967]

19. Free C, Knight R, Robertson S, Whittaker R, Edwards P, Zhou W, et al. Smoking cessation support delivered via mobile phone text messaging (txt2stop): a single-blind, randomised trial. Lancet 2011 Jul 2;378(9785):49-55 [FREE Full text] [doi: 10.1016/S0140-6736(11)60701-0] [Medline: 21722952]

20. Keoleian V, Polcin D, Galloway G. Text messaging for addiction: a review. J Psychoactive Drugs 2015;47(2):158-176 [FREE Full text] [doi: 10.1080/02791072.2015.1009200] [Medline: 25950596]

XSL·FO
RenderX

21. Robinson S, Perkins S, Bauer S, Hammond N, Treasure J, Schmidt U. Aftercare intervention through text messaging in the treatment of bulimia nervosa--feasibility pilot. Int J Eat Disord 2006 Dec;39(8):633-638. [doi: 10.1002/eat.20272] [Medline: 16937381]

22. Suffoletto B, Kristan J, Callaway C, Kim KH, Chung T, Monti PM, et al. A text message alcohol intervention for young adult emergency department patients: a randomized clinical trial. Ann Emerg Med 2014 Dec;64(6):664-72.e4 [FREE Full text] [doi: 10.1016/j.annemergmed.2014.06.010] [Medline: 25017822]

23. Moore SC, Crompton K, van Goozen S, van den Bree M, Bunney J, Lydall E. A feasibility study of short message service text messaging as a surveillance tool for alcohol consumption and vehicle for interventions in university students. BMC Public Health 2013;13:1011 [FREE Full text] [doi: 10.1186/1471-2458-13-1011] [Medline: 24160674]

24. Irvine L, Falconer DW, Jones C, Ricketts IW, Williams B, Crombie IK. Can text messages reach the parts other process measures cannot reach: an evaluation of a behavior change intervention delivered by mobile phone? PLoS One 2012;7(12):e52621 [FREE Full text] [doi: 10.1371/journal.pone.0052621] [Medline: 23300726]

25. Gonzales R, Ang A, Murphy D, Glik D, Anglin M. Substance use recovery outcomes among a cohort of youth participating in a mobile-based texting aftercare pilot program. J Subst Abuse Treat 2014 Jul;47(1):20-26 [FREE Full text] [doi: 10.1016/j.jsat.2014.01.010] [Medline: 24629885]

26. Keoleian V, Stalcup SA, Polcin DL, Brown M, Galloway G. A cognitive behavioral therapy-based text messaging intervention for methamphetamine dependence. J Psychoactive Drugs 2013;45(5):434-442 [FREE Full text] [doi: 10.1080/02791072.2013.847995] [Medline: 24592670]

27. Kuntsche E, Robert B. Short message service (SMS) technology in alcohol research--a feasibility study. Alcohol Alcohol 2009;44(4):423-428 [FREE Full text] [doi: 10.1093/alcalc/agp033] [Medline: 19482879]

28. Suffoletto B, Callaway C, Kristan J, Kraemer K, Clark D. Text-message-based drinking assessments and brief interventions for young adults discharged from the emergency department. Alcohol Clin Exp Res 2012 Mar;36(3):552-560. [doi: 10.1111/j.1530-0277.2011.01646.x] [Medline: 22168137]

29. Rios-Bedoya CF, Hay C. Feasibility of using text messaging for unhealthy behaviors screening in a clinical setting: a case study on adolescent hazardous alcohol use. J Am Med Inform Assoc 2013;20(2):373-376 [FREE Full text] [doi: 10.1136/amiajnl-2011-000688] [Medline: 22759622]

30. Bendtsen M, Bendtsen P. Feasibility and user perception of a fully automated push-based multiple-session alcohol intervention for university students: randomized controlled trial. JMIR Mhealth Uhealth 2014;2(2):e30 [FREE Full text] [doi: 10.2196/mhealth.3233] [Medline: 25098296]

31. Gonzales R, Douglas AM, Glik DC. Exploring the feasibility of text messaging to support substance abuse recovery among youth in treatment. Health Educ Res 2014 Feb;29(1):13-22 [FREE Full text] [doi: 10.1093/her/cyt094] [Medline: 24038196]

32. Haug S, Schaub MP, Venzin V, Meyer C, John U, Gmel G. A pre-post study on the appropriateness and effectiveness of a web-and text messaging-based intervention to reduce problem drinking in emerging adults. J Med Internet Res 2013;15(9):e196 [FREE Full text] [doi: 10.2196/jmir.2755] [Medline: 23999406]

33. Stoner SA, Hendershot CS. A randomized trial evaluating an mHealth system to monitor and enhance adherence to pharmacotherapy for alcohol use disorders. Addict Sci Clin Pract 2012;7:9 [FREE Full text] [doi: 10.1186/1940-0640-7-9] [Medline: 23186301]

34. Montes JM, Medina E, Gomez-Beneyto M, Maurino J. A short message service (SMS)-based strategy for enhancing adherence to antipsychotic medication in schizophrenia. Psychiatry Res 2012 Dec 30;200(2-3):89-95. [doi: 10.1016/j.psychres.2012.07.034] [Medline: 22901437]

35. Pijnenborg GH, Withaar FK, Brouwer WH, Timmerman ME, van den Bosch RJ, Evans JJ. The efficacy of SMS text messages to compensate for the effects of cognitive impairments in schizophrenia. Br J Clin Psychol 2010 Jun;49(Pt 2):259-274. [doi: 10.1348/014466509X467828] [Medline: 19735607]

36. Välimäki M, Hätönen H, Adams CE. Mobile.net: mobile telephone text messages to encourage adherence to medication and to follow up with people with psychosis: methods and protocol for a multicenter randomized controlled two-armed trial. JMIR Res Protoc 2012;1(2):e8 [FREE Full text] [doi: 10.2196/resprot.2136] [Medline: 23611874]

37. Ben-Zeev D, Kaiser SM, Krzos I. Remote "hovering" with individuals with psychotic disorders and substance use: feasibility, engagement, and therapeutic alliance with a text-messaging mobile interventionist. J Dual Diagn 2014;10(4):197-203 [FREE Full text] [doi: 10.1080/15504263.2014.962336] [Medline: 25391277]

38. Depp C, Mausbach B, Granholm E, Cardenas V, Ben-Zeev D, Patterson T, et al. Mobile interventions for severe mental illness: design and preliminary data from three approaches. J Nerv Ment Dis 2010 Oct;198(10):715-721 [FREE Full text] [doi: 10.1097/NMD.0b013e3181f49ea3] [Medline: 20921861]

39. Ainsworth J, Palmier-Claus JE, Machin M, Barrowclough C, Dunn G, Rogers A, et al. A comparison of two delivery modalities of a mobile phone-based assessment for serious mental illness: native smartphone application vs text-messaging only implementations. J Med Internet Res 2013;15(4):e60 [FREE Full text] [doi: 10.2196/jmir.2328] [Medline: 23563184]

40. Granholm E, Ben-Zeev D, Link P, Bradshaw K, Holden J. Mobile Assessment and Treatment for Schizophrenia (MATS): a pilot trial of an interactive text-messaging intervention for medication adherence, socialization, and auditory hallucinations. Schizophr Bull 2012 May;38(3):414-425 [FREE Full text] [doi: 10.1093/schbul/sbr155] [Medline: 22080492]

41.   van den Berg N, Grabe HJ, Freyberger HJ, Hoffmann W. A telephone-and text-message based telemedical care concept for patients with mental health disorders--study protocol for a randomized, controlled study design. BMC Psychiatry 2011;11:30 [FREE Full text] [doi: 10.1186/1471-244X-11-30] [Medline: 21329513]

42.   Bopp JM, Miklowitz DJ, Goodwin GM, Stevens W, Rendell JM, Geddes JR. The longitudinal course of bipolar disorder as revealed through weekly text messaging: a feasibility study. Bipolar Disord 2010 May;12(3):327-334 [FREE Full text] [doi: 10.1111/j.1399-5618.2010.00807.x] [Medline: 20565440]

43.   Agyapong VI, Ahern S, McLoughlin DM, Farren CK. Supportive text messaging for depression and comorbid alcohol use disorder: single-blind randomised trial. J Affect Disord 2012 Dec 10;141(2-3):168-176. [doi: 10.1016/j.jad.2012.02.040] [Medline: 22464008]

44.   Agyapong VI, Milnes J, McLoughlin DM, Farren CK. Perception of patients with alcohol use disorder and comorbid depression about the usefulness of supportive text messages. Technol Health Care 2013;21(1):31-39. [doi: 10.3233/THC-120707] [Medline: 23358057]

45.   Aguilera A, Berridge C. Qualitative feedback from a text messaging intervention for depression: benefits, drawbacks, and cultural differences. JMIR Mhealth Uhealth 2014;2(4):e46 [FREE Full text] [doi: 10.2196/mhealth.3660] [Medline: 25373390]

46.   Aguilera A, Muñoz RF. Text messaging as an adjunct to CBT in low-income populations: a usability and feasibility pilot study. Prof Psychol Res Pr 2011 Dec 1;42(6):472-478 [FREE Full text] [doi: 10.1037/a0025499] [Medline: 25525292]

47.   Agyapong VI, McLoughlin DM, Farren CK. Six-months outcomes of a randomised trial of supportive text messaging for depression and comorbid alcohol use disorder. J Affect Disord 2013 Oct;151(1):100-104. [doi: 10.1016/j.jad.2013.05.058] [Medline: 23800443]

48.   Shapiro J, Astin J, Shapiro SL, Robitshek D, Shapiro DH. Coping with loss of control in the practice of medicine. Fam Syst Health 2011 Mar;29(1):15-28. [doi: 10.1037/a0022921] [Medline: 21417521]

49.   Bauer S, Okon E, Meermann R, Kordy H. Technology-enhanced maintenance of treatment gains in eating disorders: efficacy of an intervention delivered via text messaging. J Consult Clin Psychol 2012 Aug;80(4):700-706. [doi: 10.1037/a0028030] [Medline: 22545736]

50.   Berrouiguet S, Alavi Z, Vaiva G, Courtet P, Baca-García E, Vidailhet P, et al. SIAM (suicide intervention assisted by messages): the development of a post-acute crisis text messaging outreach for suicide prevention. BMC Psychiatry 2014;14:294 [FREE Full text] [doi: 10.1186/s12888-014-0294-8] [Medline: 25404215]

51.   Furber G, Jones GM, Healey D, Bidargaddi N. A comparison between phone-based psychotherapy with and without text messaging support in between sessions for crisis patients. J Med Internet Res 2014;16(10):e219 [FREE Full text] [doi: 10.2196/jmir.3096] [Medline: 25295667]

52.   Price M, Ruggiero KJ, Ferguson PL, Patel SK, Treiber F, Couillard D, et al. A feasibility pilot study on the use of text messages to track PTSD symptoms after a traumatic injury. Gen Hosp Psychiatry 2014;36(3):249-254 [FREE Full text] [doi: 10.1016/j.genhosppsych.2014.02.004] [Medline: 24636721]

53.   Kunigiri G, Gajebasia N, Sallah D. Improving attendance in psychiatric outpatient clinics by using reminders. J Telemed Telecare 2014 Dec;20(8):464-467. [doi: 10.1177/1357633X14555642] [Medline: 25316039]

54.   Branson CE, Clemmey P, Mukherjee P. Text message reminders to improve outpatient therapy attendance among adolescents: a pilot study. Psychol Serv 2013 Aug;10(3):298-303. [doi: 10.1037/a0026693] [Medline: 23937089]

55.   Chandra PS, Sowmya HR, Mehrotra S, Duggal M. 'SMS' for mental health—feasibility and acceptability of using text messages for mental health promotion among young women from urban low income settings in India. Asian J Psychiatr 2014 Oct;11:59-64. [doi: 10.1016/j.ajp.2014.06.008] [Medline: 25453699]

56.   Liu Q, Abba K, Alejandria MM, Sinclair D, Balanag VM, Lansang MA. Reminder systems to improve patient adherence to tuberculosis clinic appointments for diagnosis and treatment. Cochrane Database Syst Rev 2014;11:CD006594 [FREE Full text] [doi: 10.1002/14651858.CD006594.pub3] [Medline: 25403701]

57.   Berrouiguet S, Gravey M, Le Galudec M, Alavi Z, Walter M. Post-acute crisis text messaging outreach for suicide prevention: a pilot study. Psychiatry Res 2014 Jul 30;217(3):154-157. [doi: 10.1016/j.psychres.2014.02.034] [Medline: 24736112]

58.   Stockwell MS, Kharbanda EO, Martinez RA, Vargas CY, Vawdrey DK, Camargo S. Effect of a text messaging intervention on influenza vaccination in an urban, low-income pediatric and adolescent population: a randomized controlled trial. JAMA 2012 Apr 25;307(16):1702-1708. [doi: 10.1001/jama.2012.502] [Medline: 22535855]

59.   Tan WL, Asahar SF, Harun NL. Insulin therapy refusal among type II diabetes mellitus patients in Kubang Pasu district, the state of Kedah, Malaysia. Singapore Med J 2015 Apr;56(4):224-227 [FREE Full text] [doi: 10.11622/smedj.2014170] [Medline: 25532511]

60.   Chiang T, Wang C, Lin Y, Chou S, Wang C, Juang H, et al. Factors related to treatment refusal in Taiwanese cancer patients. Asian Pac J Cancer Prev 2015;16(8):3153-3157 [FREE Full text] [Medline: 25921113]

61.   Pantell M, Rehkopf D, Jutte D, Syme SL, Balmes J, Adler N. Social isolation: a predictor of mortality comparable to traditional clinical risk factors. Am J Public Health 2013 Nov;103(11):2056-2062. [doi: 10.2105/AJPH.2013.301261] [Medline: 24028260]

62.   Yesufu-Udechuku A, Harrison B, Mayo-Wilson E, Young N, Woodhams P, Shiers D, et al. Interventions to improve the experience of caring for people with severe mental illness: systematic review and meta-analysis. Br J Psychiatry 2015 Apr;206(4):268-274 [FREE Full text] [doi: 10.1192/bjp.bp.114.147561] [Medline: 25833867]
63.   Nurmatov U, Buckingham S, Kendall M, Murray SA, White P, Sheikh A, et al. Effectiveness of holistic interventions for people with severe chronic obstructive pulmonary disease: systematic review of controlled clinical trials. PLoS One 2012;7(10):e46433 [FREE Full text] [doi: 10.1371/journal.pone.0046433] [Medline: 23110052]
64.   Hale K, Capra S, Bauer J. A framework to assist health professionals in recommending high-quality apps for supporting chronic disease self-management: illustrative assessment of type 2 diabetes apps. JMIR Mhealth Uhealth 2015;3(3):e87 [FREE Full text] [doi: 10.2196/mhealth.4532] [Medline: 26369346]
65.   Odeny TA, Newman M, Bukusi EA, McClelland RS, Cohen CR, Camlin CS. Developing content for a mHealth intervention to promote postpartum retention in prevention of mother-to-child HIV transmission programs and early infant diagnosis of HIV: a qualitative study. PLoS One 2014;9(9):e106383 [FREE Full text] [doi: 10.1371/journal.pone.0106383] [Medline: 25181408]
66.   Motto JA, Bostrom AG. A randomized controlled trial of postcrisis suicide prevention. Psychiatr Serv 2001 Jun;52(6):828-833. [Medline: 11376235]
67.   Pew Research Center. 2015 Apr 9. Teens, social media & technology overview 2015 URL: http://www.pewinternet.org/files/2015/04/PI_TeensandTech_Update2015_0409151.pdf [accessed 2015-06-03] [WebCite Cache ID 6atZcDPxl]
68.   LeRouge C, Van Slyke C, Seale D, Wright K. Baby boomers' adoption of consumer health technologies: survey on readiness and barriers. J Med Internet Res 2014;16(9):e200 [FREE Full text] [doi: 10.2196/jmir.3049] [Medline: 25199475]
69.   Garcia C, Hardeman RR, Kwon G, Lando-King E, Zhang L, Genis T, et al. Teenagers and texting: use of a youth ecological momentary assessment system in trajectory health research with latina adolescents. JMIR Mhealth Uhealth 2014;2(1):e3 [FREE Full text] [doi: 10.2196/mhealth.2576] [Medline: 25098355]
70.   Person AK, Blain ML, Jiang H, Rasmussen PW, Stout JE. Text messaging for enhancement of testing and treatment for tuberculosis, human immunodeficiency virus, and syphilis: a survey of attitudes toward cellular phones and healthcare. Telemed J E Health 2011 Apr;17(3):189-195 [FREE Full text] [doi: 10.1089/tmj.2010.0164] [Medline: 21457085]
71.   Mbuagbaw L, van der Kop ML, Lester RT, Thirumurthy H, Pop-Eleches C, Smieja M, et al. Mobile phone text messages for improving adherence to antiretroviral therapy (ART): a protocol for an individual patient data meta-analysis of randomised trials. BMJ Open 2013;3(5):e002954 [FREE Full text] [doi: 10.1136/bmjopen-2013-002954] [Medline: 23793704]
72.   Pop-Eleches C, Thirumurthy H, Habyarimana J, Zivin J, Goldstein M, de Walque D, et al. Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: a randomized controlled trial of text message reminders. AIDS 2011 Mar 27;25(6):825-834 [FREE Full text] [doi: 10.1097/QAD.0b013e32834380c1] [Medline: 21252632]

Abbreviations

EMA: ecological momentary assessment
mHealth: mobile-health
PTSD: post-traumatic stress disorder
RCT: randomized control trial
SMS: short message service

XSL·FO
RenderX

XSL•FO

RenderX

Original Paper

A Web-Based and Mobile Health Social Support Intervention to Promote Adherence to Inhaled Asthma Medications: Randomized Controlled Trial

Justin T Koufopoulos[1], MSc; Mark T Conner[1], PhD; Peter H Gardner[1], PhD; Ian Kellar[1], DPhil

School of Psychology, Faculty of Medicine and Health, University of Leeds, Leeds, United Kingdom

Corresponding Author:
Ian Kellar, DPhil
School of Psychology
Faculty of Medicine and Health
University of Leeds
Lifton Place
Leeds, LS2 9JT
United Kingdom
Phone: 44 (0)1133437242
Fax: 44 (0)1133435749
Email: i.kellar@leeds.ac.uk

Abstract

Background: Online communities hold great potential as interventions for health, particularly for the management of chronic illness. The social support that online communities can provide has been associated with positive treatment outcomes, including medication adherence. There are few studies that have attempted to assess whether membership of an online community improves health outcomes using rigorous designs.

Objective: Our objective was to conduct a rigorous proof-of-concept randomized controlled trial of an online community intervention for improving adherence to asthma medicine.

Methods: This 9-week intervention included a sample of asthmatic adults from the United Kingdom who were prescribed an inhaled corticosteroid preventer. Participants were recruited via email and randomized to either an "online community" or "no online community" (diary) condition. After each instance of preventer use, participants (N=216) were required to report the number of doses of medication taken in a short post. Those randomized to the online community condition (n=99) could read the posts of other community members, reply, and create their own posts. Participants randomized to the no online community condition (n=117) also posted their medication use, but could not read others' posts. The main outcome measures were self-reported medication adherence at baseline and follow-up (9 weeks postbaseline) and an objective measure of adherence to the intervention (visits to site).

Results: In all, 103 participants completed the study (intervention: 37.8%, 39/99; control: 62.2%, 64/117). MANCOVA of self-reported adherence to asthma preventer medicine at follow-up was not significantly different between conditions in either intention-to-treat (P=.92) or per-protocol (P=.68) analysis. Site use was generally higher in the control compared to intervention conditions.

Conclusions: Joining an online community did not improve adherence to preventer medication for asthma patients. Without the encouragement of greater community support or more components to sustain engagement over time, the current findings do not support the use of an online community to improve adherence.

ClinicalTrial: International Standard Randomized Controlled Trial Number (ISRCTN): 29399269; http://www.isrctn.com/ISRCTN29399269/29399269 (Archived by WebCite at http://www.webcitation.org/6fUbEuVoT)

(J Med Internet Res 2016;18(6):e122) doi:10.2196/jmir.4963

XSL·FO
RenderX

Introduction

Online communities have the potential to foster feelings of social support in patients battling chronic health issues [1-3]. The management of chronic health issues falls mostly on the patient and their family and can create a sense of isolation and distress [4]. Although there is no single accepted definition of an online community, a common definition is "...a group of people who share a strong common interest, form relationships, and interact online" [1]. Some of the largest online communities for patients (as of November 2015) are MedHelp, PatientsLikeMe, TuDiabetes, and DailyStrength with millions of members [5]. Such online communities can provide valuable support, but do they also relate to better management of chronic health issues such as increased medication adherence?

Link Between Social Support and Medication Adherence

Medication adherence can be defined as the extent to which a patient follows medication-taking guidelines agreed on by a patient and doctor [6]. Common guidelines include medication dosage and frequency. According to the World Health Organization, adherence to asthma medicine is just 50%, representing a significant health threat including increased risk of hospitalization and death [4]. Adherence is a complex phenomenon influenced by many factors, including economic (eg, financial costs of drugs or therapies), social (eg, age, race), therapy (eg, adverse reactions to medication), and patient (eg, individual attitudes and concerns) factors [4,6].

Adherence to medication is also influenced by social support. Although many of the factors relating to adherence are difficult to influence, social support is a promising target because it may be improved through low cost interventions (eg, support groups, sponsors). In a systematic review of 122 studies published between 1948 and 2001, DiMatteo [7] found a significant relationship between social support and adherence. Studies were categorized into types of support, including practical support (eg, instrumental support, assistance, reminders, organization, support for a specific behavior), emotional support, unidimensional social support (involving multiple types of social support, not separated in their measurement), family cohesiveness (eg, warmth, closeness, acceptance), and marital status and living arrangement. Patients receiving practical support were 3.6 times more likely to adhere to treatment regimens than those who were not. Risk of nonadherence was also found to be 1.53 times more likely if patients had low social support. Equivalence between online and face-to-face interventions in related fields suggests that online support could similarly provide social support and be associated with improved medication adherence [8].

Current State of the Evidence for Online Communities: Randomized Controlled Trials

There have been only a few randomized controlled trials (RCTs) of online communities for patients with chronic health issues [9-13]. The results of these studies have been mixed. For example, a randomized controlled trial by Richardson et al [9], found an online community for an Internet-mediated walking

program did not increase participant step count, but participants randomized to the online community had greater engagement and lower rates of attrition than the control group.

Similarly, a RCT by Brindal et al [10] of an online platform for weight loss found that compared to the noninteractive control group, groups with online community features (eg, friend requests, newsfeeds, quizzes, profile pages) had greater engagement, but did not show increased weight loss or retention. A RCT by Stoddard et al [11] also found that an online community feature (a message board) appeared to increase engagement for a smoking cessation website, but this feature did not influence quit rates.

Although more broadly fitting the definition of an online community [1], we found two trials that used Facebook to evaluate the effectiveness of online communities on physical activity, and weight loss, respectively [12,13]. In a RCT to evaluate the feasibility and efficacy of a Facebook-based intervention on physical activity for young adult cancer survivors, participants were randomized to either a Facebook online community intervention condition (FITNET) or a Facebook self-care condition. Increases in light physical activity were more than 2 hours per week greater in the FITNET condition compared to the self-care condition [12]. In another trial using Facebook, students with a body mass index of 25-50 kg/m^2 were randomized to Facebook, a Facebook plus text messaging and personal feedback group, or a wait list. After 8 weeks, the Facebook plus text messaging and personal feedback group had significantly greater weight loss than either of the other two groups [13].

Theoretical Framework

We predicted that participating in an online community would lead to greater medication adherence. The theoretical underpinnings of this prediction are Social Cognitive Theory [14,15], the Theory of Planned Behavior (TPB) [16], and the stress and coping perspective of social support [17]. The effect of community website exposure [11] is also included in the theoretical framework of the intervention. According to Social Cognitive Theory [14,15], individuals can learn by observing the actions of others. If those actions produce an effect that is beneficial to the individual being observed, those actions are more likely to be imitated. It was predicted that participants observing the adherence of other patients will themselves improve adherence to inhaled corticosteroid (ICS) treatment. Participants reading other patients' success stories regarding adherence or dealing with asthma more generally will learn from these stories and apply these lessons to their own life, improving adherence.

The TPB states that the more an individual intends to perform a given behavior, the more likely they are to perform that behavior [16]. According to the TPB, intentions to perform a behavior are based on norms, attitudes, perceived behavioral control, and influence whether a given behavior is performed through intentions. It is predicted that when patients observe self-reports of adherence and other interactions on the site, over time these observations will positively influence their attitudes, norms, and perceived behavioral control around adherence

promoting stronger intentions to adhere and subsequent adherence behavior.

Social support has been defined as the quality and structure of an individual's relationships; greater social support is associated with improvements in adherence to medication regimens [7]. As is predicted in the stress-buffering perspective of social support [17], the perception of having socially supportive relationships and the support that participants actually receive will reduce stress associated with adherence and asthma, improving adherence and overall health.

Objective and Hypothesis

Our objective was to conduct a rigorous study of the effects of participating in an online community on adherence to asthma medicine. To our knowledge, there are no previous studies of the relationship between online communities and medication adherence. Asthma was chosen as the target illness for this trial because medication adherence is often low and the incidence of chronic asthma in adults is relatively high: nearly 10% in the United Kingdom [18]. Asthma in adults is typically treated with a combination of an ICS preventer and a bronchodilator reliever. We hypothesized that adherence would be improved by participation in an online community due to processes of modeling and social support.

Methods

Study Design

In this 2-arm RCT, participants were enrolled at random into either the intervention condition, "AsthmaVillage," an online community for patients with asthma, or the control condition, "AsthmaDiary," an online diary for recording ICS preventer use. Intervention arm participants had access to an online community and could leave comments or see who else was online. In contrast, the control-arm participants could not read the posts of other control-arm participants or interact with other participants online. An active control was used to test the effect of the community on adherence and to prevent participants from guessing if they were in the group of interest. The study was carried out for 9 weeks, between June 24 and August 26, 2013. The trial conformed to the Consolidated Standards of Reporting Trials (CONSORT) eHealth Checklist (Multimedia Appendix 1) [19].

Recruitment

A total of 1833 emails requesting participants for a study on asthma management were sent out to department secretaries of the 40 largest universities in the United Kingdom by enrollment. Universities were chosen as recruitment sites because recruitment through medical centers or primary care practices in the United Kingdom requires a lengthy approvals process that can go on for many months, which was beyond the resources available for this study.

Recruitment emails were sent over a period of 10 days from June 13 to 23, 2013. Department secretaries were asked in the body of the email to forward the request to department mailing lists (Multimedia Appendix 2). The request for participants invited individuals managing their asthma with an ICS preventer to fill out an eligibility screening form and included a link to the questionnaire (Multimedia Appendix 3). Participants were also informed that on successful completion of the study they would receive a £20 (approximately US $30) shopping voucher. Successful completion of the study was defined as recording their ICS preventer use at least once per week on the site for the duration of the study.

Eligibility Screening and Consent

A total of 936 participants responded to the eligibility questionnaire. Participants were excluded from the study if they failed to complete the eligibility questionnaire (n=256) or baseline measures (n=228), did not have asthma (n=105), were not prescribed an ICS preventer inhaler for a weekly regimen of at least one dose per week (n=87), failed to complete informed consent (n=35), or had previously participated in the pilot study (n=9). After screening, a total of 251 participants were eligible for study inclusion. See Figure 1 for details.

Participants were automatically taken to the information sheet (Multimedia Appendix 4) and were asked to provide informed consent (Multimedia Appendix 5). In all, 35 participants refused to provide consent and were eliminated from the study, leaving 216 eligible participants. Participants were then randomized to the diary (n=117) and online community (n=99) conditions. Randomization occurred through a random number generator [20], yielding two unequal groups. The experimenters then manually separated the two lists and emailed both groups log-in instructions.

The online screening survey was administered through Qualtrics [21], a subscription-based online survey software suite.

Figure 1. Participant flowchart of recruitment, participation after exclusion criteria, randomization, and attrition.
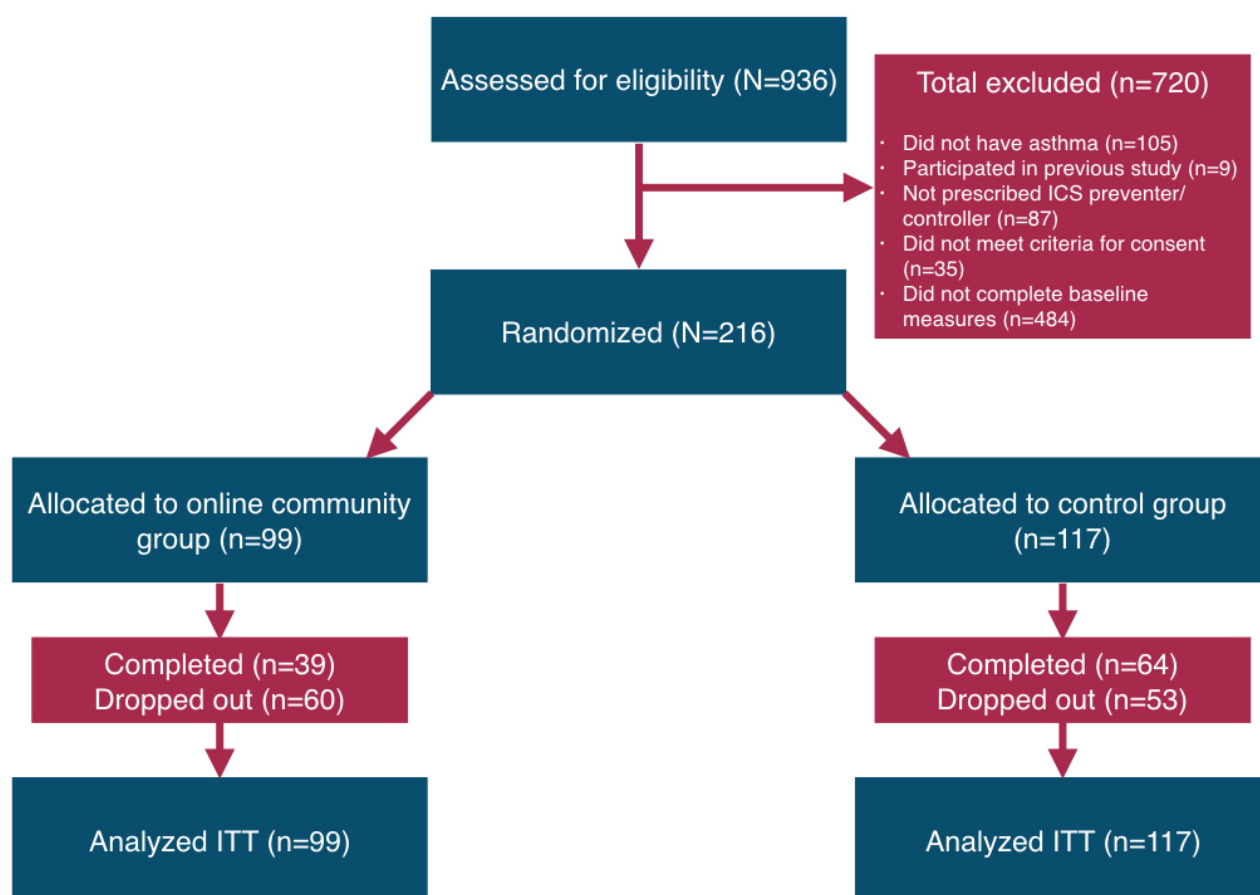
Baseline Measures

Baseline measures were included as part of the eligibility screening. Participants completed an online survey (Multimedia Appendix 6) that included questions about gender, age, previous social networking use, and prescriptions.

Preventer adherence in both conditions was self-reported using the 6-item Simplified Medication Adherence Questionnaire (SMAQ) [22] (Multimedia Appendix 7), a common, validated measure of medication adherence. A self-report measure of medication adherence was used because this study was designed to be low cost, feasible, and widely geographically distributed in the United Kingdom. Furthermore, questionnaires can have a high concordance with more expensive objective measures, such as electronic counters [23]. In the questionnaire, the SMAQ refers generally to all medicine. For example, the first item of the SMAQ is "Do you ever forget to take your medicine?" For this study, all instances of the word "medicine" were changed to "asthma preventer medication." This small change was unlikely to have affected the measure.

The SMAQ was then recalculated with dichotomous scoring of all variables (more than two missed uses was treated as nonadherent) and by reverse scoring of item 4 of the SMAQ ("Thinking about the last week, how often have you not taken your asthma preventer medicine as prescribed?"). The rescored SMAQ formed a reliable measure (Cronbach alpha=.72);

therefore, a mean across all completed items was computed. If there were two or fewer completed SMAQ items for an individual, the score was treated as missing (ie, no value was given) and these individuals were excluded from the analysis involving this variable. Replacing missing values with imputed values did not substantively alter the reported findings.

Pilot Study

Previous to this trial, a small pilot study (N=8) was conducted to gather qualitative feedback on the usability of the online community. The online community was created using WordPress [24], an open-source content management system, and BuddyPress [25] social networking features (Multimedia Appendix 8). There were three primary intervention components: (1) a home page that displayed all site activity in a rolling status board, (2) a group diary for posting preventer use, and (3) a profile page. The intervention was developed and adapted based on feedback collected during the pilot study. No additional feedback was collected during the trial.

The pilot study ran for 31 days. Participants were required to post their preventer use weekly in the online community. At the end of the study, participants were sent a questionnaire with short-answer items regarding the usefulness and usability of the site. Participants reported that they found the social support features of the intervention useful in connecting with other asthma patients, but found that after the first few weeks it was

difficult to become engaged with the site because the number of active conversations diminished.

The results of the pilot study influenced the development of the final intervention administered during the RCT. The main finding from the pilot study was that participants found it difficult to engage with the site because of the lack of activity. With only eight members, participant conversations fell off rapidly and little discussion was observed after the first week. Participants reported logging in to the website to report their medication usage and then logging out right after with little else happening on the site to engage them. However, most of the participants reported that hearing about other asthma sufferers' experiences and having a forum to ask questions about asthma was useful. To more directly engage participants, a separate site section dedicated to questions about asthma was created for the RCT.

Additionally, participants reported often failing to remember to log in to AsthmaVillage, which likely also affected site usage. Therefore, for the RCT we decided to implement a system of automated, weekly reminder notifications. This automation was accomplished with MailChimp, an email marketing software tool.

Intervention

Like the pilot, the online community was created using WordPress and BuddyPress. In addition to the three features mentioned in relation to the pilot study, a fourth feature was added based on the results of the pilot and the need to increase engagement: a page for posting questions and answers about asthma.

The main actions participants could take on the online community were reporting their preventer use and writing posts, comments, or questions. Questions and comments needed to be answered by the community members themselves because there was no experimenter intervention once the trial had begun. The only feedback patients could receive during the trial was from other patients themselves because this intervention was optimized for implementation at scale and at low cost. This trial attempted to understand the value of an online community, implemented without the added support of a community manager to engage members. The intervention was developed to be accessible by smartphone Web browsers as well as by desktop versions.

The effect of membership in the online community is dependent on the extent members use the website of the online community itself. As previously mentioned, site engagement was a barrier for participants in the pilot study. In order to create a website that was more engaging, we created a site discussion section dedicated to posting questions and answers regarding asthma. Such features have been shown to be beneficial to engagement in previous studies [9].

Control

The control condition comprised an online diary, AsthmaDiary. The online diary was created using Google Forms. A single-item survey was created (Multimedia Appendix 9): "How many times did you take your preventer?" Participants randomized to the control condition could then input the number of puffs and, after entering their unique personal identification number (PIN), hit "submit." Because participants did not need to log in with a username to fill out the form, participants used a PIN that allowed their posts to be identified by the researcher. Participants in the control condition could not see the posts of the other participants or otherwise know that there were other participants posting in their condition.

Follow-Up Measures

Follow-up measures were taken 9 weeks postbaseline. The SMAQ was also used to measure ICS preventer adherence at follow-up (SMAQ-T2) and scored in an identical fashion. The SMAQ-T2 formed a reliable measure (Cronbach alpha=.69) and mean scores were computed across all completed items. Missing values were assigned to any participant with fewer than two filled-out SMAQ-T2 items. Intention-to-treat (ITT) analyses used SMAQ-T1 means carried forward to follow-up for participants who did not complete the follow-up SMAQ-T2 measures.

Site activity was measured by producing counts of comments and posts by user and week. Comments were exported from the WordPress content management system and categorized into the following groupings: "preventer posts" (posts that were self-reports about preventer use), "posts about symptoms" (posts made by users about their asthma symptoms), "questions about asthma" (questions about asthma embedded in a comment or post, or standalone), "answer or reply comments" (answers or comments left by users to posted questions), and "nonanswer comments" (comments or statements left by users that were neither about symptoms or supporting another user). Posts or comments were categorized at the sentence level because longer posts or comments often had statements about symptoms and follow-up questions in the same post.

Participants in both conditions were instructed to report their preventer use on their assigned website each time they used their preventers over a period of 9 weeks. The extent that participants adhered to these directions was calculated (site adherence). The total number of preventer puffs (total puffs) each participant reported taking in each week was divided by the total number of preventer puffs prescribed in a given week (daily total number of puffs prescribed in 1 day multiplied by 7). This calculation produced a score of weekly site adherence for each participant. Thus, if a participant logged their preventer use every time they were prescribed to use their preventer, their site adherence would be 100%. The mean adherence across each of the 9 weeks was calculated, forming an overall score of adherence to the site (site adherence).

Procedure

After randomization, participants were emailed instructions on how to use their site (Multimedia Appendices 10 and 11). Participants were then prompted to log in to their respective websites. Before logging in, participants created a username and password for their sites. Participants randomized to the diary created a PIN that they were required to input whenever they posted preventer use. Participants in both conditions were quasi-anonymous. Registration required an email confirmation.

It was not possible to determine whether an individual operated multiple accounts on AsthmaVillage or the online diary. Account passwords were screened by the experimenters for duplicates in an attempt to mitigate this possibility.

During the trial, an automated weekly email was sent to participants indicating which week the trial was on (eg, week 6 of 9) and with a reminder to post their preventer inhaler use (Multimedia Appendix 12). At the end of the study, participants were emailed instructions (Multimedia Appendix 13) on how to complete the follow-up measures.

Participants who completed the study and posted on their site at least once a week were mailed a £20 (approximately US $30) shopping voucher for participation (n=82). Participants who completed at least the baseline and follow-up measures were mailed a £10 (approximately US $15) voucher (n=23). This approach was taken to allow as many participants as possible an opportunity to fill out the follow-up questionnaire and reduce bias in the sample.

Statistical Analysis

Sample Size Calculation

Based on an expected medium effect size (d=.5), an alpha of .05 (1-tailed), and power of 80%, we calculated that a total of 102 participants would be needed to complete the study. A medium effect size was justified based on the review by Webb et al [26], which found that greater use of theory in Internet interventions was associated with larger effect sizes. For example, use of the TPB was associated with a medium-sized effect (d=.5) in this review [26].

In the few RCTs of online communities, attrition varied considerably. We assumed a 50% dropout rate and, thus, aimed to recruit double the number of participants the power analyses suggested were required.

Analysis

First, descriptive statistics were calculated for all variables (gender, age, total puffs, site adherence, SMAQ-T1, SMAQ-T2, and SMAQ-T2 ITT) and examined across the whole sample and for each condition to ensure the measures were normally distributed.

Next, we examined the effect of condition on outcome variables. Multivariate analysis of variance (MANOVA) was used when variables were not measured at baseline (total puffs and site adherence). Multivariate analysis of covariance (MANCOVA) was used when variables were also measured at baseline (SMAQ) with the baseline score being the covariate.

Statistical analyses were conducted using SPSS version 20.0 (IBM Corp, Armonk, NY, USA).

Human Participants and Trial Registration

The University of Leeds, School of Psychology Ethics Committee approved this study (ethics reference number 13-0096). All participants gave online consent. The details of the trial were made public in advance (ISRCTN trial registration number: 29399269).

Results

Attrition

Of the 216 participants who met our inclusion criteria, only 103 participants fully completed the study, 64 (62.1%) of these were from the control arm. Of the 99 participants allocated to the intervention arm, 82 created a username and password (83%). A chi-square test indicated that dropout was higher in the intervention condition (60/99, 61%) than the control condition (53/117, 45.3%; χ^2_2=5.0, P=.03).

We also tested whether the sample who completed the study were representative of the initial sample on baseline measures of gender, age, and SMAQ-T1. MANOVA revealed a significant difference between groups (Wilks' lambda=0.944, $F_{3,212}$=4.190, P=.007). Examination of the univariate effects revealed significant effects for SMAQ-T1 ($F_{1,214}$=4.48, P=.04) and gender ($F_{1,214}$=7.20, P=.008), but no effect for age ($F_{1,214}$=0.72). On average, completers scored higher on the SMAQ-T1 (mean 1.49, SD 0.30) than noncompleters (mean 1.40, SD 0.27). Higher SMAQ scores indicate lower preventer adherence. Study completers were also more likely to be female (79.6%, 82/103) than noncompleters (61.1%, 69/113).

Our attrition analyses indicated that the sample completing the study was not fully representative of those starting the study; therefore, our analyses based on completers should be treated with caution.

Descriptive Statistics

Participants were mostly women and although ages ranged from 18 to 64 years, the average participant was in their late twenties (Table 1). There was generally an even mixture of adherent and nonadherent participants. Table 1 also shows that, in general, the control and intervention arms showed few differences except in relation to site adherence measures.

Table 1. Descriptive statistics.

Variable	Total N=216	Control n=117	Intervention n=99
Age (years), mean (SD)	28.1 (9.7)	28.8 (10.1)	27.2 (9.2)
Gender, n (%)			
Male	64 (29.6)	35 (29.9)	29 (29.3)
Female	151 (69.9)	82 (70.1)	69 (69.7)
Self-report adherence, mean (SD)			
SMAQ-T1	1.45 (0.29)	1.48 (0.30)	1.41 (0.27)
SMAQ-T2[a]	1.48 (0.28)	1.49 (0.29)	1.46 (0.28)
SMAQ-T2 ITT	1.44 (0.28)	1.46 (0.29)	1.42 (0.26)
Site measures, mean (SD)			
Site adherence[b]	18.35 (21.98)	25.53 (25.63)	9.22 (10.87)
Total puffs[c]	42.32 (50.40)	54.49 (58.01)	21.20 (20.80)

[a]For SMAQ-T2, total n=104, control n=64, and intervention n=40.

[b]For site adherence, total n=134, control n=75, and intervention n=59.

[c]For total puffs, total n=134, control n=85, and intervention n=49.

Primary Outcomes

MANCOVA indicated no condition effects on the SMAQ scores (Wilks' lambda=0.998; $F_{1,96}$=0.176, P=.68). For this analysis, 104 participants were included, with 64 in the control and 40 in the intervention groups. When based on ITT analyses, MANCOVA also indicated no condition effects for the SMAQ, with 117 participants in the control and 99 participants in the intervention groups (Wilks' lambda=1.000; $F_{1,207}$=0.011 , P=.92). MANOVA for variables not measured at baseline (site adherence, total puffs) revealed significant differences for condition (Wilks' lambda=0.922; $F_{2,114}$=4.835, P=.01). Examination of the univariate statistics revealed significant differences for both site adherence ($F_{1,91}$=6.635, P=.01) and total puffs ($F_{1,91}$=9.400, P=.003). Examination of the means revealed that these differences reflected the higher levels of site adherence and total puffs in the control compared to the intervention condition (Table 1).

Site Activity Via Community Posts and Comments

Examination of site activity via the posts and comments created by users during the 9 weeks suggested that the majority of site activity were self-reports of asthma preventer use (Figure 2). Of the 99 participants allocated to the intervention condition, 83 created a username and password.

Of the 861 comments or posts left by users on AsthmaVillage (intervention condition), 754 (87.6%) were preventer posts (eg, "2x preventer" and "1x symbiocort"). Looking at the remaining comments or posts, 7.1% (61/754) were posts about symptoms (eg, "Got my preventer back, asthma has been quite uncontrolled this past week" and "No need to take my ventolin today, feeling good " and "Well still haven't taken my preventer since my last post. To be honest, not feeling any different"), 1.9% (16/754) were questions (eg, "Has anyone taken Singulair? I've heard it's good as a preventer and for allergies. I've been on

fexofenadine for years and it's starting to lose effectiveness. It would be great to take a tablet that would replace my antihistamine and my symbicort" and "I have a question about asthma reviews at the doctor. They chase me to come every 6 months, but nothing has ever changed [I'm 35 and have had asthma since I was 5, so I'm pretty good at managing it myself by now]. I don't really understand the point of the review—if my symptoms get worse I will go to the doctor, but if it's well managed why do I still need to go every 6 months? Does anyone else find these reviews useful? Is there something I can do to get more out of them?" and "I'm self-regulating my dosage at the moment and my symptoms are almost nonexistent during summer. I was wondering if it's better to keep taking the preventer two puffs a day, or to stop the preventer and just use reliever before exercise? Anyone know?"), 1.7% (15/754) were answer comments (eg, "Hi, I have the same experiences. Nothing has changed in 20 years and it feels like a waste of their time. I'm keen to know if there's anything I can do to get more out of them too." and "Since I've been with my current GP I've had only annual reviews [and they are useful in that the preventer was altered from a pure steroid to include a long acting broncho-dilator that's rendered Ventolin seldom used]."), and 1.7% (15/754) were nonanswer statements (eg, "Finally managed to pick up my preventer inhaler today after far too long without it." and "I've got a cold...").

In all, 33 of 82 participants (40%) posted something on the site that was not purely a preventer post (ie, posts about symptoms, questions about asthma, answer or reply comments, nonanswer comments), and there were a mean 3.24 (SD 0.94) nonpreventer posts over the 9 weeks of the intervention. Eight of 82 participants (10%) explicitly asked questions of the community and tended to post more frequently (mean number of nonpreventer posts in this subgroup was 5.38, SD 3.50).

XSL•FO
RenderX

Site Adherence Over Nine Weeks

Further detailed examination of the weekly site adherence means over the 9 weeks of the intervention indicated a substantial difference between conditions in adherence. This was particularly apparent at week 1, with 41.7% adherence in the control condition compared to 11.3% site adherence in the intervention condition. Figure 3 shows that site adherence was most different between groups at week 1, but fell at a much greater rate across weeks in the control compared to the intervention condition. Site adherence was relatively consistent across weeks in the intervention condition.

Figure 2. Comments and posts by type over nine weeks.
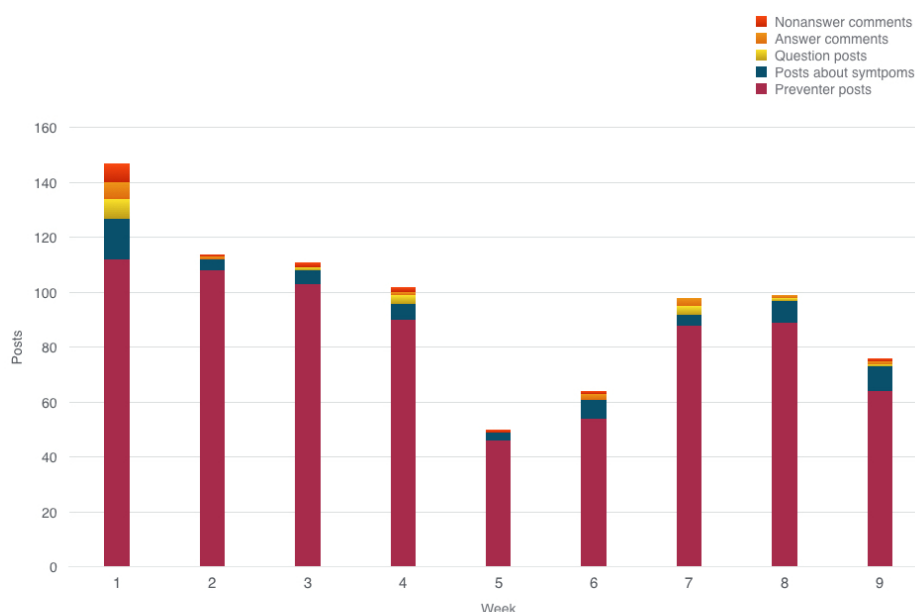
Figure 3. Change in site adherence over nine weeks by condition.

XSL•FO
RenderX
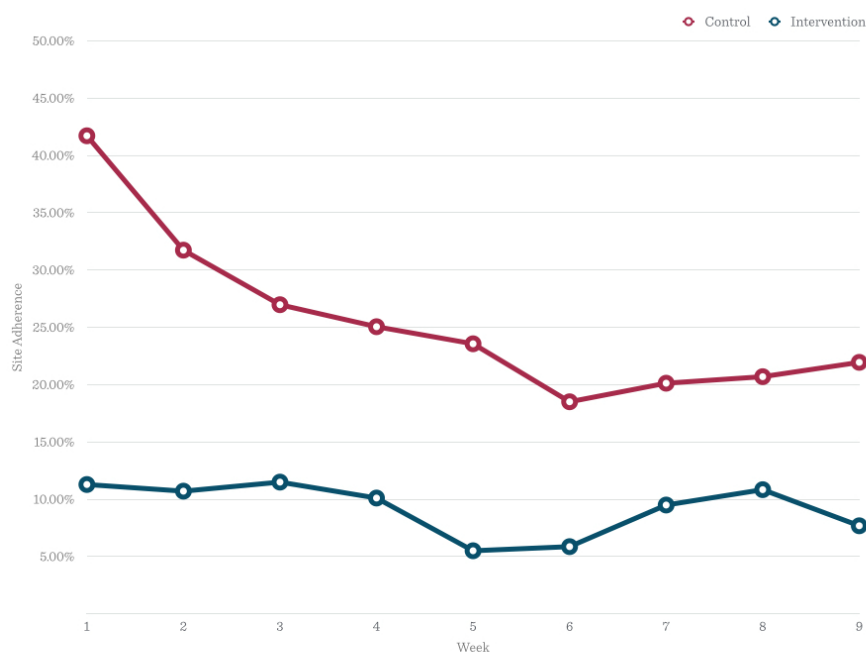
Discussion

Summary of Principal Results

This RCT examined whether being part of an online community would improve self-reported preventer adherence. We anticipated that through the mechanisms of role modeling, social support, and website exposure the intervention compared to the control condition would increase adherence. However, contrary to expectations, being part of an online community for asthma patients for 9 weeks failed to increase self-reported medication adherence to ICS preventer therapy compared to a control diary condition (P=.68). In addition, there was significantly lower site adherence in the intervention condition than in the control condition (P<.001), even from the first week of the intervention (Figure 3).

Study condition also predicted attrition. Participants were less likely to complete the study if they were randomized to the intervention compared to control condition (P=.03). Further examination of this data indicated that both baseline adherence as assessed by SMAQ scores (P=.04) and gender (P=.008) were related to attrition with study completers being more likely to be women (79.6%, 82/103 vs 61.1%, 69/113) and less likely to be adherent to asthma preventer medication at baseline than noncompleters (mean 1.49, SD 0.30 vs mean 1.41, SD 0.27). This finding can perhaps be explained by the site being more useful for people struggling with asthma preventer adherence and less so for people without problems. Such an interpretation would be consistent with Magnezi et al [3] and their study of patient activation or the extent individuals are able to manage their own health care. These authors found a negative relationship between patient activation and perceived usefulness of a website because taking a less active role in one's own medical care predicted higher website usefulness.

Recruitment and Attrition

Of the 936 participants who began the eligibility and baseline screening forms, 48.3% (452/936) completed them and, after exclusions, 23.1% (216/936) were randomized to condition. It is not unusual for Web-based studies of online communities to have a large number of participants dropout between screening and randomization. For example, Richardson et al [9] had 880 signups, but after dropouts and exclusions, only 324 were randomized. It is also possible that given the number of question items and forms that needed to be completed to be eligible for the study, participants who were initially attracted by the monetary reward became discouraged and failed to continue.

Of the 216 participants who began the trial, 103 (47.7%) completed the 9-week study. Attrition was within the typical range described by other online community RCTs [9-13] and reviews of the literature for midsized trials [26]. However, it is difficult to make comparisons because these RCTs are so different in design and virtually none tested the effectiveness of the online community as a standalone intervention as in this study.

Primary Outcomes

The online community did not improve self-reported preventer adherence (SMAQ) compared to the control. This finding is consistent with findings by Eysenbach et al [27] that virtual health communities were not associated with improved health outcomes and, more recently, Richardson et al [9] or Brindal et al [10] that found membership in online communities had no effect on behavior change. The evidence from these trials and the one reported here would suggest that joining an online community intervention is not associated with improved health behavior.

Site Activity Via Community Posts and Comments

Analysis of the comments and posts indicates that the online community was primarily used for posting asthma medication use (Figure 2). This is not entirely unexpected because posting ICS preventer use was the only requirement for the study. However, more than 100 (12%) of the 754 posts were nonpreventer and 40% (33/83) of the community could be considered as "nonlurkers." Within the context of online health communities, "lurkers" are individuals who do not participate in posting [28]. Because these members posted something other than a preventer post, which was a study requirement, these members met the criteria for nonlurkers. Studies indicate lurking to be highly variable, between 0 and 99% [29]. Nonnecke and Preece [29] found a mean 45.5% of lurkers in online health communities. Research indicates that both lurkers and nonlurkers can receive benefits from online health communities [30]. Overall, the present online community did foster interaction between participants (eg, questions, answers to questions, posts about symptoms) and there existed a reasonable ratio of lurkers to nonlurkers.

Site Adherence Over Nine Weeks

Beginning in the first week of the study, there was a significant difference between conditions for site adherence. One possible explanation could be that participants did not like posting their preventer use in an online community compared to the participants posting in an online diary. Perhaps these participants felt worried or uncomfortable posting this information publicly even though their identity was anonymized. On the one hand, an online diary could maintain a sense of privacy; on the other hand, the more rapid decline in posting in the diary condition might be explained by a lack of engagement over time. Engagement remained fairly consistent in the online community perhaps because of the presence of other members. Such an explanation would also be consistent with the findings of Richardson et al [9], in which an online community was found to reduce attrition to an Internet-mediated walking program, but did not increase walking step count.

Study Strengths and Weaknesses

This study had a number of strengths and weaknesses. In relation to strengths, this study tested the effect of membership in an online community as a single intervention component in a RCT; RCTs are the gold standard for determining the effect of an intervention on an outcome. Second, previous studies have also attempted to influence the member activity of online communities through various time-intensive posting strategies, possibly confounding their results [9,27]. In contrast, this study did not attempt to influence participation beyond the weekly reminders sent to participants in both conditions. Outside of a

research context, it is improbable that organizations seeking to enhance health behaviors would divert considerable resources toward encouraging participation. As such, this study represents a test of a deliverable intervention. Third, the inclusion criteria for participation were broad, primarily requiring that patients be prescribed an ICS preventer for daily use. This increases the generalizability of the findings to a large percentage of individuals with asthma.

In relation to weaknesses, the only validated measure of adherence was a self-report measure. A self-report measure was chosen to facilitate the aims of this study, which was to deliver an intervention at low cost across a geographically widespread group of participants in the United Kingdom. Several studies have shown that self-report measures of adherence can be unreliable [31,32]. However, in an analysis of 86 published studies using both self-report and non–self-report measures, Garber et al [23] found that self-report questionnaires and diaries were the most highly concordant with electronic measures (75% agreement).

Second, previous RCTs of online communities [9-13] have involved multiple components beyond the community itself to influence health behavior change. As such, these studies are unable to isolate the effectiveness of a single component compared to another, making it difficult to determine causality of the reported effects. Although it was a deliberate choice to test only a single component for this study, reliance on a single component also made this intervention more susceptible to failure. Similar to the pilot study, it may be that when the online community was unable to foster sustained engagement, participation dropped off.

Third, the measure site adherence should also be interpreted with caution. It is not an objective measure of actual preventer adherence, but more likely a measure of adherence to the study reporting instructions. It is possible that actual preventer adherence was different from that reported by participants on either site. Such an explanation would be supported by the SMAQ-T2 ITT and SMAQ-T2 means (Table 1), which indicated that there were no significant differences in self-report adherence for either condition at follow-up.

Finally, because universities were chosen as recruitment sites, it is possible that individuals associated with universities (eg,

teachers, students, staff) are more likely to have the type of economic and social stability that may be associated with better adherence, self-reported or otherwise. This may affect generalizability because the most poorly controlled asthma patients may not be well represented in our sample.

Future Directions

The present findings indicate that a "pure" online community does not improve self-report medication adherence. Future research may wish to experiment with multiple levels of engagement with an online community. For example, researchers may have a pure community in the comparator and a community with a virtual coach in the intervention condition. Automation and machine learning could also be used predict compliance and offer different levels of automated support, such as notifications on relevant content or reminders to take medication. These types of community interventions might be better tailored to individual participants with varying levels of adherence or different attitudes around asthma and compliance. Although we believe assessing the independent value of an online community component is an important step for online health community research, without enough participants to sustain long-term discussion, pure community interventions are likely to fail. Thus, additional patient-tailored components can offer more value to patients and possibly sustain engagement over longer periods of time.

Because self-report measures of adherence are sometimes unreliable, future studies may wish to also invest in a mechanical or digital measure of adherence able to provide objective ICS preventer data that can be compared against self-report measures such as those employed here.

Conclusions

An online community did not improve self-reported adherence to asthma preventer medicine. Surprisingly, participants were much more adherent in posting inhaler use in the control condition than the intervention condition, although it appeared that this difference between conditions attenuated over time. Without greater community support beyond the existence of the community itself, it does not seem that an online community alone can improve adherence. However, our analyses of attrition suggest that online communities may be more useful to patients with poor asthma adherence than patients with good adherence.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH (V 1.6.1)-Submission/Publication Form.

XSL•FO

RenderX

[PDF File (Adobe PDF File), 1MB - jmir_v18i6e122_app1.pdf]

Multimedia Appendix 2

Recruitment email.

[PDF File (Adobe PDF File), 71KB - jmir_v18i6e122_app2.pdf]

Multimedia Appendix 3

Eligibility screening form.

[PDF File (Adobe PDF File), 36KB - jmir_v18i6e122_app3.pdf]

Multimedia Appendix 4

Participant information sheet.

[PDF File (Adobe PDF File), 72KB - jmir_v18i6e122_app4.pdf]

Multimedia Appendix 5

Participant Consent Form.

[PDF File (Adobe PDF File), 46KB - jmir_v18i6e122_app5.pdf]

Multimedia Appendix 6

Baseline questionnaire.

[PDF File (Adobe PDF File), 63KB - jmir_v18i6e122_app6.pdf]

Multimedia Appendix 7

Simplified Medication Adherence Questionnaire.

[PDF File (Adobe PDF File), 44KB - jmir_v18i6e122_app7.pdf]

Multimedia Appendix 8

Screenshots of website homepage, group diary page, asthma Q&A page, and profile page.

[PDF File (Adobe PDF File), 370KB - jmir_v18i6e122_app8.pdf]

Multimedia Appendix 9

Control condition diary screenshot.

[PNG File, 28KB - jmir_v18i6e122_app9.png]

Multimedia Appendix 10

Participant site instructions.

[PDF File (Adobe PDF File), 152KB - jmir_v18i6e122_app10.pdf]

Multimedia Appendix 11

Participant site instructions.

[PDF File (Adobe PDF File), 321KB - jmir_v18i6e122_app11.pdf]

Multimedia Appendix 12

Reminder to post preventer inhaler.

[PDF File (Adobe PDF File), 62KB - jmir_v18i6e122_app12.pdf]

Multimedia Appendix 13

Instructions on how to complete the follow-up measures.

[PDF File (Adobe PDF File), 78KB - jmir_v18i6e122_app13.pdf]

References

1. Young C. Community management that works: how to build and sustain a thriving online health community. J Med Internet Res 2013;15(6):e119 [FREE Full text] [doi: 10.2196/jmir.2501] [Medline: 23759312]
2. van der Eijk M, Faber MJ, Aarts JW, Kremer JA, Munneke M, Bloem BR. Using online health communities to deliver patient-centered care to people with chronic conditions. J Med Internet Res 2013 Jun;15(6):e115 [FREE Full text] [doi: 10.2196/jmir.2476] [Medline: 23803284]
3. Magnezi R, Bergman YS, Grosberg D. Online activity and participation in treatment affects the perceived efficacy of social health networks among patients with chronic illness. J Med Internet Res 2014;16(1):e12 [FREE Full text] [doi: 10.2196/jmir.2630] [Medline: 24413148]
4. World Health Organization. 2003. Adherence to long-term therapies: evidence for action URL: http://apps.who.int/iris/bitstream/10665/42682/1/9241545992.pdf [accessed 2016-05-05] [WebCite Cache ID 6hHcB3ASV]
5. Swan M. Crowdsourced health research studies: an important emerging complement to clinical trials in the public health research ecosystem. J Med Internet Res 2012;14(2):e46 [FREE Full text] [doi: 10.2196/jmir.1988] [Medline: 22397809]
6. Horne R, Weinman J, Barber N, Elliott R, Morgan M. Concordance, Adherence and Compliance in Medicine Taking: Report for the National Co-ordinating Centre for NHS Service Delivery and Organisation R & D. London: National Coordinating Centre for the Service Delivery and Organisation; 2005 Dec. URL: http://www.netscc.ac.uk/hsdr/files/project/SDO_FR_08-1412-076_V01.pdf [accessed 2016-05-05] [WebCite Cache ID 6hHcISs3F]
7. DiMatteo MR. Social support and patient adherence to medical treatment: a meta-analysis. Health Psychol 2004 Mar;23(2):207-218. [doi: 10.1037/0278-6133.23.2.207] [Medline: 15008666]
8. van Ballegooijen W, Cuijpers P, van Straten A, Karyotaki E, Andersson G, Smit JH, et al. Adherence to Internet-based and face-to-face cognitive behavioural therapy for depression: a meta-analysis. PLoS One 2014;9(7):e100674 [FREE Full text] [doi: 10.1371/journal.pone.0100674] [Medline: 25029507]
9. 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]
10. Brindal E, Freyne J, Saunders I, Berkovsky S, Smith G, Noakes M. Features predicting weight loss in overweight or obese participants in a web-based intervention: randomized trial. J Med Internet Res 2012;14(6):e173 [FREE Full text] [doi: 10.2196/jmir.2156] [Medline: 23234759]
11. Stoddard JL, Augustson EM, Moser RP. Effect of adding a virtual community (bulletin board) to smokefree.gov: randomized controlled trial. J Med Internet Res 2008;10(5):e53 [FREE Full text] [doi: 10.2196/jmir.1124] [Medline: 19097974]
12. Valle CG, Tate DF, Mayer DK, Allicock M, Cai J. A randomized trial of a Facebook-based physical activity intervention for young adult cancer survivors. J Cancer Surviv 2013 Sep;7(3):355-368 [FREE Full text] [doi: 10.1007/s11764-013-0279-5] [Medline: 23532799]
13. Napolitano MA, Hayes S, Bennett GG, Ives AK, Foster GD. Using Facebook and text messaging to deliver a weight loss program to college students. Obesity (Silver Spring) 2013 Jan;21(1):25-31 [FREE Full text] [doi: 10.1002/oby.20232] [Medline: 23505165]
14. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [Medline: 847061]
15. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs, NJ: Prentice-Hall; 1986.
16. Ajzen I. The Theory of Planned Behavior. Organ Behav Hum Decis Process 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
17. Lakey B, Cohen S. Social support theory measurement. In: Cohen S, Underwood LG, Gottlieb BH, editors. Social Support Measurement and Intervention: A Guide for Health and Social Scientists. New York: Oxford University Press; 2000:29-52.
18. NHS. Asthma URL: http://www.nhs.uk/Conditions/Asthma/Pages/Introduction.aspx [accessed 2016-05-05] [WebCite Cache ID 6hHcyxDST]
19. 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]
20. RANDOM.ORG. URL: https://www.random.org/ [accessed 2016-05-05] [WebCite Cache ID 6hHeFnLBs]
21. Qualtrics. URL: https://www.qualtrics.com/ [accessed 2016-05-05] [WebCite Cache ID 6hHeOZTkC]
22. Knobel H, Alonso J, Casado JL, Collazos J, González J, Ruiz I, et al. Validation of a simplified medication adherence questionnaire in a large cohort of HIV-infected patients: the GEEMA Study. AIDS 2002 Mar 8;16(4):605-613. [Medline: 11873004]

XSL•FO

RenderX

23. Garber MC, Nau DP, Erickson SR, Aikens JE, Lawrence JB. The concordance of self-report with other measures of medication adherence: a summary of the literature. Med Care 2004 Jul;42(7):649-652. [Medline: 15213489]

24. WordPress. URL: https://wordpress.com/ [accessed 2015-02-07] [WebCite Cache ID 6WAkrOb1d]

25. BuddyPress. URL: https://buddypress.org/ [accessed 2016-05-05] [WebCite Cache ID 6hHeVY0Ic]

26. 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]

27. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. BMJ 2004 May 15;328(7449):1166 [FREE Full text] [doi: 10.1136/bmj.328.7449.1166] [Medline: 15142921]

28. van Uden-Kraan CF, Drossaert CH, Taal E, Seydel ER, van de Laar MA. Self-reported differences in empowerment between lurkers and posters in online patient support groups. J Med Internet Res 2008;10(2):e18 [FREE Full text] [doi: 10.2196/jmir.992] [Medline: 18653442]

29. Nonnecke B, Preece J. Lurker demographics: counting the silent. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2000 Presented at: CHI '00 SIGCHI Conference on Human Factors in Computing Systems; Apr 1-6, 2000; The Hague, The Netherlands p. 73-80 URL: http://www.cis.uoguelph.ca/~nonnecke/research/demographics.pdf

30. Setoyama Y, Yamazaki Y, Namayama K. Benefits of peer support in online Japanese breast cancer communities: differences between lurkers and posters. J Med Internet Res 2011;13(4):e122 [FREE Full text] [doi: 10.2196/jmir.1696] [Medline: 22204869]

31. Garber MC, Nau DP, Erickson SR, Aikens JE, Lawrence JB. The concordance of self-report with other measures of medication adherence: a summary of the literature. Med Care 2004 Jul;42(7):649-652. [Medline: 15213489]

32. Shi L, Liu J, Fonseca V, Walker P, Kalsekar A, Pawaskar M. Correlation between adherence rates measured by MEMS and self-reported questionnaires: a meta-analysis. Health Qual Life Outcomes 2010;8:99 [FREE Full text] [doi: 10.1186/1477-7525-8-99] [Medline: 20836888]

Abbreviations

ICS: inhaled corticosteroid
ITT: intention-to-treat
MANOVA: multivariate analysis of variance
PIN: personal identification number
RCT: randomized controlled trial
SMAQ: Simplified Medication Adherence Questionnaire
TPB: Theory of Planned Behavior

XSL•FO
RenderX

Original Paper

A Technology-Mediated Behavioral Weight Gain Prevention Intervention for College Students: Controlled, Quasi-Experimental Study

Delia Smith West[1], PhD; Courtney M Monroe[1], Ph. D.; Gabrielle Turner-McGrievy[1], PhD, RD; Beth Sundstrom[2], PhD; Chelsea Larsen[1], M.P.H., CHES; Karen Magradey[1]; Sara Wilcox[1], PhD; Heather M Brandt[1], PhD

[1]Arnold School of Public Health, University of South Carolina, Columbia, SC, United States

[2]Department of Communication, College of Charleston, Charleston, SC, United States

Corresponding Author:
Delia Smith West, PhD
Arnold School of Public Health
University of South Carolina
915 Greene St
Discovery Building I, Suite 403
Columbia, SC, 29802
United States
Phone: 1 803 777 8826
Fax: 1 803 777 5457
Email: westds@mailbox.sc.edu

Abstract

Background: Both men and women are vulnerable to weight gain during the college years, and this phenomenon is linked to an increased risk of several chronic diseases and mortality. Technology represents an attractive medium for the delivery of weight control interventions focused on college students, given its reach and appeal among this population. However, few technology-mediated weight gain prevention interventions have been evaluated for college students.

Objective: This study examined a new technology-based, social media-facilitated weight gain prevention intervention for college students.

Methods: Undergraduates ($n = 58$) in two sections of a public university course were allocated to either a behavioral weight gain prevention intervention (Healthy Weight, HW; N=29) or a human papillomavirus (HPV) vaccination awareness intervention (control; N=29). All students were enrolled, regardless of initial body weight or expressed interest in weight management. The interventions delivered 8 lessons via electronic newsletters and Facebook postings over 9 weeks, which were designed to foster social support and introduce relevant educational content. The HW intervention targeted behavioral strategies to prevent weight gain and provided participants with a Wi-Fi-enabled scale and an electronic physical activity tracker to facilitate weight regulation. A repeated-measures analysis of variance was conducted to examine within- and between-group differences in measures of self-reported weight control practices and objectively measured weight. Use of each intervention medium and device was objectively tracked, and intervention satisfaction measures were obtained.

Results: Students remained weight stable (HW: −0.48+1.9 kg; control: −0.45+1.4 kg), with no significant difference between groups over 9 weeks ($P = .94$). However, HW students reported a significantly greater increase in the number of appropriate weight control strategies than did controls (2.1+4.5 vs −1.1+3.4, respectively; $P = .003$) and there was no increase in inappropriate weight control behaviors ($P = .11$). More than 90% of students in the HW arm opened the electronic newsletters each week, and the average number of Facebook interactions (comments and likes) per student each week was 3.3+1.4. Each self-monitoring device was initialized by 90% of HW students. On average, they used their physical activity tracker for 23.7+15.2 days and their Wi-Fi scale for 14.1+13.1 days over the 9 weeks. HW students rated the intervention favorably.

Conclusions: The short-term effect of this technology-based weight gain prevention intervention for college students is promising and merits evaluation over a longer duration to determine whether engagement and behavioral improvements positively affect weight outcomes and can be maintained.

Introduction

Young adulthood represents a period in which weight gain and the onset of obesity are common [1,2], and these patterns are associated with an increased risk of chronic diseases and mortality [3-5]. The phenomenon of weight gain among college freshmen in particular is well known, with one recent meta-analysis reporting increases in weight ranging from 0.73 to 3.99 kg and an average weight gain of 1.74 kg among studies which objectively measured students' weight [6]. Furthermore, weight gain continues throughout the full college period, and both men and women are vulnerable to this pattern [7-10]. Of note, an estimated one-third of college students are overweight or obese [11]. These facts are especially concerning given that college represents a critical transition when young adults begin to make independent decisions about a range of personal choices, including lifestyle behaviors [1]. The physical activity and dietary patterns they adopt—the two key lifestyle factors influencing weight [12]—will likely track into adulthood [1]. It is evident that the college years present a prime target period in which effective weight management should be promoted in an effort to shape the future health and well-being of a large number of individuals.

Technology offers an attractive platform for behavioral weight control interventions targeting college students because it is both familiar and appealing to young adults. The vast majority of college students regularly use the Internet (99%) [13] and mobile phones (80%) [14] to access information and connect socially, particularly through social networking platforms. Most college students (80%) use online social networks [13]; Facebook is the most widely used social networking platform among individuals aged between 18 and 29 years [15]. Furthermore, advanced technologies, such as "wearables" (eg, electronic physical activity trackers) and electronically-enabled health monitoring devices (eg, Wi-Fi body weight scales), are increasingly available and hold promise for managing weight [16-18]. These technologies provide opportunities for both collecting objective data and delivering intervention components in real-time, bringing interventions to the setting in which behaviors occur.

Interventions to promote healthy diet [19-23], physical activity [24-28], and weight control [29-33] among college students have shown potential. However, only a small number of studies have evaluated the effectiveness of technology-based interventions designed specifically for weight gain prevention among college students [34-36], and the findings from these studies have been mixed. There is a clear need for the continued development and evaluation of technology-based weight gain prevention interventions targeting college students. Indeed, in a recent review of the limited body of research on weight gain prevention interventions in college students, Laska et al [37] called for the further design and evaluation of interventions that capitalize on the crucial influences of technology and social networks for this population. Importantly, no healthy weight interventions targeting college students have sought to harness the joint capabilities of advanced technological monitoring devices and online social networks. The purpose of this study was to examine a novel, technology-mediated weight gain prevention intervention for college students.

Methods

Study Design and Procedures

This controlled, quasi-experimental study recruited undergraduates enrolled in two sections of an advanced health communication class at a public university in the Southeastern United States. Baseline questionnaires were administered via a secure website (Qualtrics, Provo, Utah), and anthropometric data were collected in-person by the study staff. Both classes received an 8-session, technology-mediated health promotion intervention that provided health education and facilitated social support, delivered over 9 weeks. Classes were randomly allocated by a coin toss to either (1) a behavioral weight gain prevention intervention (Healthy Weight, HW) or to (2) a human papillomavirus (HPV) vaccination awareness intervention (control). Each health promotion intervention served as a control for the other. This study design allowed for the simultaneous implementation and evaluation of interventions centered on two areas of health promotion that are particularly relevant to college students.

The HW intervention focused on weight gain patterns among college populations and long-term risk for obesity-related chronic diseases, as well as behavioral strategies to maintain weight and avoid obesity. Participants in the HW intervention were given a Wi-Fi-enabled scale (Aria; Fitbit Inc., Boston, MA) and an electronic physical activity tracker (Fitbit Zip; Fitbit Inc., Boston, MA) to facilitate weight regulation. Participants in both interventions were emailed weekly electronic newsletters targeting relevant content and enrolled in separate, private Facebook groups. These groups served as an additional channel for the delivery of intervention content, as well as promoted interaction between fellow group members, facilitated by study counselors. The control condition did not include any healthy weight-related behavior change elements. The interventions were matched for intervention duration and structure, as well as the number of newsletters and planned Facebook content postings. Newsletters were sent automatically via email using MailChimp (Rocket Science Group, LLC, Atlanta, GA), which allows prescheduling of distribution and also tracks who opens the newsletter link; both groups received newsletters on the same schedule. Posttreatment questionnaire data and body weight measurements were obtained 9 weeks after intervention initiation. Detailed descriptions of specific measures administered are given below. Primary outcomes for the HW intervention described here were change in body weight and in self-reported weight control behaviors. The Institutional Review Boards at both the College of Charleston and University of South Carolina approved the study.

XSL•FO
RenderX

Participant Recruitment and Eligibility

Email invitations outlining the study and eligibility criteria were sent to students enrolled in the targeted courses. Each student could voluntarily elect to either participate in the study as a course assignment or engage in an alternative course activity. To be included, students were required to be registered for the course, have access to the Internet via a computer and/or mobile device, and be willing to use their existing email address and Facebook account, or establish a new email address and Facebook account, for the study. Students were also required to provide informed consent on a secure website.

Healthy Weight Intervention

The HW intervention was based on the social cognitive theory (SCT) [38] and content was adapted from the Diabetes Prevention Program Lifestyle Intervention [39] and technology-mediated behavioral weight control programs [18,40,41]. The content was delivered by weekly electronic newsletters and Facebook postings coupled with technological tools that provided objective data on physical activity and body weight. Participants were encouraged to weigh themselves daily using the Wi-Fi scale provided and to track their weight over time using the affiliated website or mobile app. The rationale for daily weighing was provided and a self-regulation approach to promote weight maintenance and stability by initiating suitable weight management behaviors when body weight increased [18,42] was advocated. Daily self-weighing has been shown to facilitate weight loss and maintenance [43,44], with no evidence of harms associated with the practice [45,46]. Participants were given methods to identify their current weight status and information that directed those who might be overweight to engage in more focused weight management efforts, whereas those who were not overweight were instructed to focus their efforts on making their dietary intake and physical activity levels healthier while maintaining their current weight.

A key strategy for promoting healthy weight and fostering positive lifestyle habits was to emphasize increased physical activity. The electronic physical activity tracker provided personalized and real-time objective feedback on steps taken and miles walked, as well as cumulative personal reports on activity level, via the associated website and mobile app. Participants were given graded goals to increase their steps to at least 10,000/day [47] and suggestions were offered about how to use the feedback from the tracker to set proximal goals and stay motivated. Campus walking routes and university gym hours were provided in lesson materials to facilitate adoption of increased physical activity.

Healthy dietary intake patterns were addressed in the intervention, and behavioral strategies for improving diet quality were offered. Facebook posts and newsletters focused on building social support for selecting more fruits and vegetables, lower-fat dairy products, and lower-calorie and/or more nutrient-dense snack options, as well as validating a social norm around eating a healthier diet. However, no specific calorie goals or food patterns were prescribed.

Lesson content focused on themes relevant to college students engaged in healthy weight regulation (Table 1). Each lesson was tailored specifically for the campus, with examples of locations or events that promoted physical activity and tips on healthy food choices in the dining halls and local restaurants. In addition, study personnel posted at least five posts each week to the private Facebook message board. Posts were scheduled to allow automatized distribution using a social media management tool (Hootsuite Media Inc., Vancouver, BC, Canada). Study investigators moderated the Facebook page, providing answers to questions or stimulating interaction, as appropriate. The posts were designed to cultivate a social climate of support for behavior change and included questions designed to prompt discussion about behavior change efforts, encourage social support from peers for new health behaviors, model successful weight control behaviors, and elicit suggestions to overcome barriers encountered. Polls were used to provide social norming for healthy behaviors. These types of messages have been shown to stimulate discussion more effectively during social media–delivered weight loss interventions [40]. Intervention activities (reading the newsletters and interacting on Facebook) were intended to take approximately 30 min/week in total. The program spanned a 9-week period during the spring semester, inclusive of spring break.

Table 1. Lesson topics delivered via electronic newsletters and Facebook posts.

Session	Topics
1	Introduction to Healthy Weight intervention, setting up monitoring devices, and on-campus resources
2	Energy balance, self-monitoring, goal setting, and overcoming barriers
3	Nutrition and healthy dietary intake practices
4	Physical activity benefits and recommendations
5	Healthy eating and physical activity specific to the university setting
6	Sugar sweetened beverages, dining out, stress, and alcohol
7	Relapse prevention and social support
8	Overview of healthy weight control practices

Control Intervention

The control group received a HPV vaccination awareness intervention with similar elements to those implemented in the HW intervention. In brief, the control intervention was intended to increase knowledge about the health consequences of HPV and the benefits of the vaccination series, which can prevent certain types of HPV. Over the 9-week period, 8 electronic newsletters, featuring content related to the risks associated with HPV and the benefits of the vaccination series, as well as identifying barriers to vaccination and strategies to overcome these barriers, were delivered. A separate, private Facebook group with five or more posts each week, that was facilitated by the research staff, was also offered to provide relevant information, to build social support for completing the full vaccination series, and to create favorable social norms around HPV prevention through vaccination. The control intervention was matched in duration and contact schedule to the HW intervention, but did not include weight management content or any information on physical activity or healthy diet, and these participants were not provided with any monitoring devices.

Measures

All measures were obtained at baseline and immediately after intervention (9 weeks after baseline) unless otherwise noted. Self-report measures were administered in an online questionnaire format with direct data entry by the participants.

Anthropometric Data

Body weight was measured to the nearest 0.1 kg in light clothing without shoes using a digital scale (Tanita BWB 800, Arlington Heights, IL). Height was measured to the nearest 0.1 cm at baseline only using a stadiometer. Body mass index (BMI) was calculated as weight (kg)/height (m^2).

Behavioral Weight Control Practices

Behavioral weight control practices were evaluated with a 28-item checklist that assessed both appropriate behavioral weight management strategies (eg, weigh yourself, record food intake, increase exercise levels; n =23) and inappropriate weight management strategies (eg, smoke cigarettes, take diet pills, avoid food for 24 hours; n =5). The inventory asks individuals to indicate whether they have engaged in the specified behaviors over the previous month and can be examined as either the total number of healthy (or unhealthy) practices endorsed or by examining specific behaviors. The appropriate or healthy weight management items were developed for use in the Look AHEAD trial, and practices identified in the measure have been associated cross-sectionally with lower adiposity [48] and were predictive of long-term weight maintenance [49]. The current study adapted the measure to include items assessing unhealthy or inappropriate weight control strategies (see Multimedia Appendix 1). Change in reported use of weight control practices from baseline to follow-up was the primary focus of the current study, with specific interest in increases in appropriate weight management practices (considered a positive change) and in inappropriate weight management behaviors (considered an iatrogenic, negative change).

Sociodemographic Information

Demographic characteristics (ie, age, sex, race, and academic year) were self-reported on the baseline questionnaire.

Intervention Engagement

Participant engagement in the HW intervention was assessed by objectively tracking interaction with each platform used and device provided. Engagement data on the newsletters were obtained from MailChimp metrics, which provide information on who opened each newsletter. If a participant opened the newsletter, he or she was considered to have engaged in the newsletter component. If the participant liked or commented on a Facebook post, he or she was considered to have engaged in that platform. In addition, the total number of Facebook likes and comments by each participant were tallied.

Fitbit Zip and Wi-Fi scale usage data were also obtained. The number of students using the devices in a given week was tracked. A participant was considered to have used the Fitbit Zip and Aria scale in a given week if at least one day of data from each respective device was available for the individual in that week. In addition to categorical data on weekly use, the absolute number of days a participant used the Aria scale and Fitbit Zip during the study period and within each week was also tracked.

Treatment Satisfaction

After the intervention, participants were asked how useful they found the program and how likely they were to recommend the program to a friend or family member; responses were recorded using a 5-point Likert-type response format. Participants were also asked to rate satisfaction with specific elements of the intervention, such as number of lessons, postings, and the devices.

Statistical Analyses

Statistical analyses were conducted using SPSS version 22.0 for Windows (IBM Corp., Armonk, NY). Descriptive statistics were calculated for all baseline and engagement measures. Baseline comparisons between conditions were made using the independent t test for continuous variables and the chi-square analysis for categorical variables. Two control participants did not complete the postintervention questionnaires (although weight data were collected postintervention for those participants). Therefore, intent-to-treat analyses (with the last value carried forward) of the numbers of appropriate and inappropriate weight control practices reported were compared between conditions. Primary outcomes (body weight and total number of behavioral weight control strategies) were examined with repeated-measures analysis of variance (with time point as the within-participant variable and group as the between-participant variable). Individual univariate comparisons between conditions for individual behavioral weight control strategies reported post-intervention were conducted using the chi-square analysis. Analysis of treatment satisfaction measures aggregated agree, somewhat agree, and strongly agree responses, and combined disagree, somewhat disagree, and strongly disagree, and descriptive data are provided. A P value of less than .05 was used to determine statistical significance.

Results

Sample Characteristics

All students in the two sections of the targeted health communication class were eligible for and elected to enroll in the study. Students (n =58) were upperclassmen who averaged 21.6±2.2 years of age and were predominantly normal weight, although the average baseline BMI was in the upper end of the normal weight range. The majority were white. No significant differences were evident in the baseline characteristics of students in the two conditions. Retention rates at the posttreatment assessment were high, with no significant difference between groups (Table 2).

Table 2. Baseline characteristics and retention rates[a].

Measure		All (n =58)	Healthy weight (n =29)	Control (n =29)	P
Age (y)		21.6 ± 2.2	22.1 ± 2.9	21.1 ± 0.8	.08
Female (%)		81	79	83	.74
White (%)		90	83	97	.05
Academic year (%)					
	Sophomore	2	3	0	
	Junior	29	14	45	
	Senior	69	83	55	.30
Weight (kg)		67.0 ± 15.8	67.3 ± 11.3	66.6 ± 19.4	.26
BMI[b] (kg/m^2)		24.0 ± 5.1	24.1 ± 4.3	23.9 ± 5.9	.37
Overweight/obese (BMI ≥ 25 BMI, %)		22	24	21	.75
Retained for follow-up (%)		97	100	93	.15

[a]Data are mean ±standard deviation unless indicated by percentage (%).

[b]BMI = body mass index.

Weight Change

Both groups remained fairly weight stable over the 9-week study period (HW: −0.48±1.9 kg; control: −0.45±1.4 kg), with no significant Group × Time interaction (P =.94). Examination of weight changes among individuals who were overweight at the start of the program (22% of sample) revealed no significant differences in weight change between conditions. Overweight students in the HW group (n =7) lost 1.8±0.7 kg after 9 weeks compared to 1.4±1.7 kg among overweight students in the control group (n =6; P =.71).

Behavioral Weight Control Practices

In contrast to weight, there was a significant Group × Time interaction with respect to the total number of appropriate weight control strategies students reported using in the previous month. An increase in the total number of these strategies was observed at postintervention for those in the HW group (2.1±4.5) versus those in the control group who experienced no significant change (−1.1±3.4; P=.003) (Table 3). Specific weight control practices that were higher at posttreatment among those in the HW group compared to those in the control group were self-weighing (P=.005), cutting out snacking (P =.001), reducing carbohydrate intake (P =.02), graphing weight (P=.01), reducing calorie intake (P =.02), reducing fat intake (P =.02), and increasing exercise (P=.02). No Group × Time effect for the total number of reported inappropriate weight control strategies was found (P=.11), and the absolute number of inappropriate strategies remained low at both time points (Table 3).

Table 3. Self-reported use of behavioral weight control strategies[a].

Measure		All (n=58)	Healthy weight (n=29)	Control (n=29)	P
Appropriate weight control strategies [b]					
Total number (mean+ SD [c] **)**					
	Pre	8.0 ± 4.1	7.8 ± 4.5	8.3 ± 3.6	
	Post	8.5 ± 3.6	9.9 ± 3.1	7.1 ± 3.6	.003[d]
Inappropriate weight control strategies [e] **, total number (mean± SD)**					
	Pre	0.3 ± 0.7	0.3 ±0.7	0.3 ± 0.6	
	Post	0.3 ± 0.6	0.4 ± 0.7	0.1 ± 0.4	.11

[a]Questionnaire available from Multimedia Appendix 1

[b]Total possible number of appropriate weight control strategies (23).

[c]SD: standard deviation.

[d]Statistically significant difference between groups at $P <.05$; Note: P values are for Group × Time interactions.

[e]Total possible number of inappropriate weight control practices (5).

Intervention Engagement and Treatment Satisfaction

Electronic newsletters were opened each week by the majority of participants in the HW condition, with all participants (n =29) opening the initial weekly newsletter and at least 26 out of 29 participants (90%) opening the newsletters during each subsequent week (Figure 1). The total number of participants who had at least one interaction on the private Facebook page (ie, liked or commented on a post) when study investigators posted ranged from 23 out of 29 (79%) during week 1 to 29 out of 29 (100%) during weeks 3 and 7 (Figure 2). Participants made a total of 862 comments and likes over the intervention period, resulting in an average of 3.3 ±1.4 per person per week, with little variation from week to week except during spring break when there were no posts provided by the intervention to prompt student response (Figure 2). In total, participants averaged 29.5 ±13.0 comments and likes over the course of the intervention, with comments representing the majority of the interactions (26.1±9.7).

Fitbit Zips were initialized by 26 out of 29 participants (90%) in the HW group, and the number of participants who used the device ranged from 24% (7 out of 29 participants) during spring break to 83% (24 out of 29 participants) during week 2, with students using the Fitbit Zip to record step counts for an average of 23.7±15.2 days across the 9-week observation period (Figure 3). The physical activity tracker was used an average of 2.6 ± 0.9 days per person per week.

Wi-Fi scales were also initialized by 26 out of 29 (90%) HW participants. Two students reported that challenges with the campus Internet were responsible for their failure to initialize. The number of participants who used their Wi-Fi scale ranged from a low of 38% (11 out of 29) during week 1 and spring break to a high of 76% (22 out of 29) during week 2 (Figure 3). On average, students used the Aria scales on 14.1±13.1 days over the intervention period or 1.6±0.7 days a week.

Overall, HW participants rated the intervention positively, with 90% (26 out of 29) indicating they enjoyed it, 86% (25 out of 29) reporting it was helpful, and 83% (24 out of 29) saying that they would recommend the program to a friend. Most participants (26 out of 29; 90%) reported that they were satisfied with the number of lessons, the number of Facebook postings (23 out of 29; 79%), the length of the Facebook postings (24 out of 29; 82%), and the extent of interaction with the study investigators on Facebook (21 out of 29; 72%). Most students also rated the electronic physical activity trackers and Wi-Fi scales positively, with 83% (24 out of 29) rating the trackers and 66% (19 out of 29) rating the Wi-Fi scales as helpful.

XSL•FO

RenderX

Figure 1. Participant engagement with healthy weight newsletters (N=29).[a] No electronic newsletters were delivered by the study investigators.
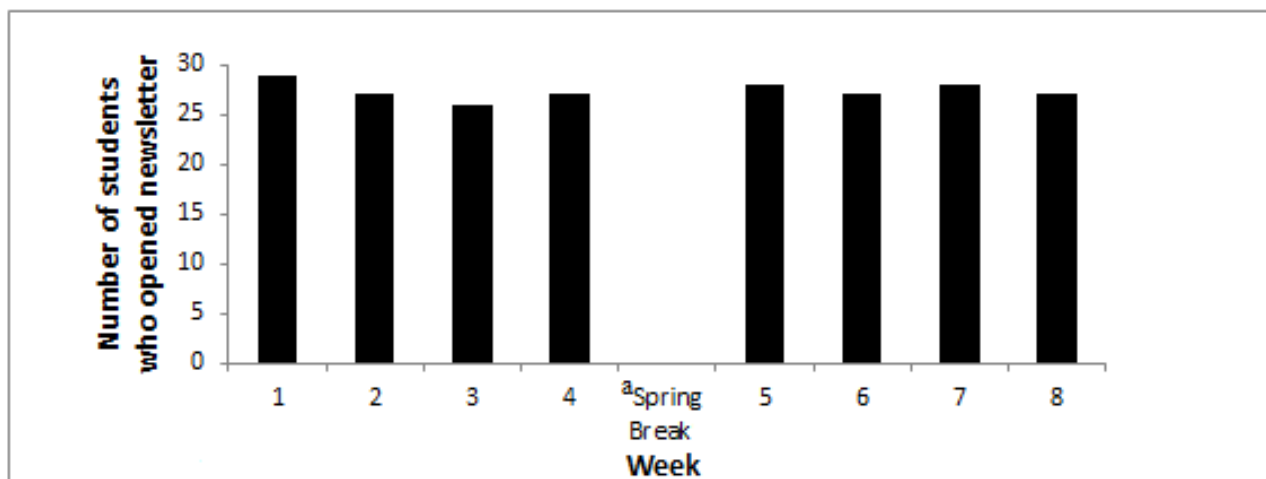
Figure 2. Participant engagement with healthy weight Facebook group (N=29).[a] No Facebook posts were made by the study investigators.
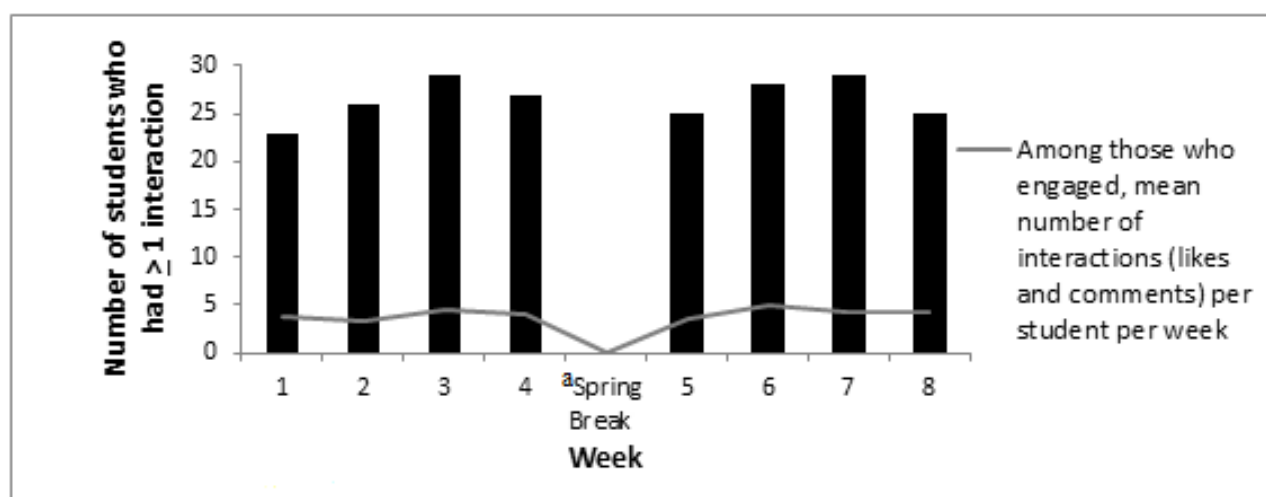
Figure 3. Participant engagement with electronic devices (N=29).

XSL•FO
RenderX
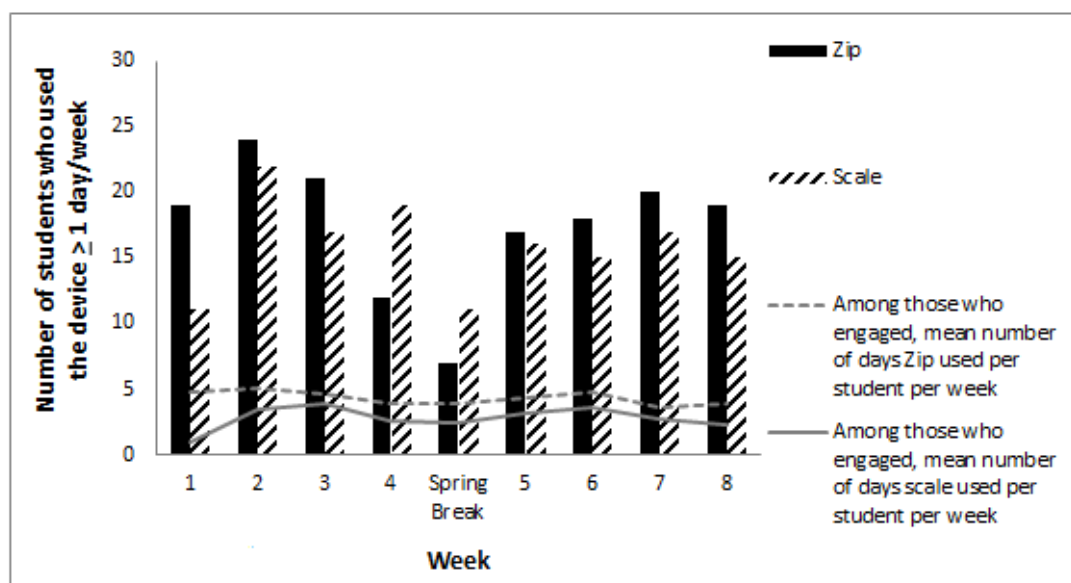
Discussion

Summary of Principal Results and Comparison to Existing Literature

A technology-mediated, theory-based weight gain prevention intervention targeted at college students that combined the use of Facebook, advanced technological monitoring devices, and electronic newsletters was well accepted by the students. Furthermore, results suggest the approach was efficacious in promoting effective weight management behaviors. The HW intervention led to a significant increase in the number of appropriate weight control behaviors reported by students receiving the intervention relative to the control group. Although no difference in body weight was apparent between conditions over the short intervention period, use of these weight control strategies over time has been shown to predict better weight management among obese individuals [49]. Therefore, continued implementation of these strategies for self-regulating weight may result in differences in weight gain over time between college students who implement more of these weight control behaviors compared with those who do not.

To our knowledge, this study is the first to use both Facebook and wireless monitoring devices to successfully deliver a Web-based weight gain prevention intervention to college students. The generally high level of engagement with the various intervention elements points to the potential this approach has for the target population. In particular, there was very high penetration of the electronic newsletter, with at least 90% of students opening the email each week. Admittedly, opening the newsletter does not indicate that students read the newsletter in detail, nor that they enacted the behavioral strategies for healthy weight management addressed in the newsletter, but it does provide an affirmative indication of treatment receipt. Furthermore, all students interacted with the Facebook group by either posting a comment or "liking" a post or comment. An average of more than three comments and likes were made per person per week, and this high participation level was sustained across the intervention period. This high degree of engagement may reflect how easily accessible these technological tools are to college students and that they are already integrated into their daily lives.

Previous studies have demonstrated that using email [50,51] and Facebook [31,51] to promote healthy lifestyle behaviors and weight loss in college students is feasible; however, the level of engagement with these features was higher in this study than in previous studies [31,50,51]. For example, one study [31] evaluated an 8-week, technology-based weight loss intervention for college students characterized in part by the use of a Facebook group. During this intervention, only 60% of participants posted at least 1 message to their Facebook group and, among those who posted, the average number of posts was approximately 2 messages. In contrast, this study had an average posting frequency of 26 postings over the course of the intervention.

The higher level of email and Facebook engagement observed in this study may be attributable, in part, to the fact that the intervention was integrated into a course for which students

received credit. Thus, they may have been more motivated to engage with these technologies than in other studies that did not embed the technology-based intervention within a course [31,50,51]. However, providing course credit or research credit as an incentive for participation in weight gain prevention studies is not uncommon [34,36,52]. Course credit alone does not assure high participation rates; several studies that provided course credit for engaging in the intervention have reported low treatment engagement and/or retention rates [34,52]. Therefore, it is possible that the intervention implemented in this study itself contributed to the high engagement and retention rates. Future weight gain prevention studies in this population should continue to evaluate the effect of providing course credit for participation on both engagement and retention over a longer observation period.

Embedding a weight gain prevention initiative in a college course may also have some particular advantages for a social network–facilitated intervention. Participants in this study may have felt a sense of connectedness as they were enrolled in the same class. In turn, this may have made them more willing to interact via Facebook. Participants in previous Facebook-based studies [31,51,53] were part of ad hoc networks, which may have made some participants reluctant to share their thoughts [54].

This study is the first to provide college students with an electronic physical activity tracker and Wi-Fi scale to promote lifestyle behaviors designed to manage weight, and data on uptake are somewhat promising. Almost all students initialized both devices, and students used the Fitbit Zip an average of 24 days or just under 3.5 weeks across the 9-week observation period. The Wi-Fi scale was used just over 1.5 days a week on average. This was lower uptake than the recommended daily use, which may reflect the fact that students were not selected based on their interest in making lifestyle behavior changes and rather were engaged in the program based on enrollment in a class. Their motivation to monitor their weight and/or their physical activity might have been higher had they volunteered for the program initially based on their interest in weight control, as in the study reported by Napolitano et al [31]. Of interest, the lowest device use was observed in this study over the spring break week. No Facebook interactions from students occurred during this period either. However, no intervention content was delivered by email or Facebook from research staff during this week; hence, there was nothing to prompt students to react. Engagement with social media and device use resumed after the break ended; however, this experience suggests that there is a need to explore ways to facilitate sustained engagement in a technology-mediated weight gain prevention intervention during academic breaks. Data suggest that academic breaks may be a particularly risky time for weight gain for younger children [55], underscoring the importance of addressing this issue. A reasonable initial step would be to continue posting Facebook messages and providing newsletters during breaks, perhaps focusing content on making healthy lifestyle choices while on vacation.

The intervention not only appears to be acceptable but also efficacious in promoting self-regulatory behaviors that have been linked to effective weight control. Noteworthy was the

absence of an increase in inappropriate weight control practices following exposure to an intervention that highlighted the need to achieve and maintain a healthy weight, a concern that has been voiced about treatments focusing on weight management among young adults [45]. The development of a pattern of healthy lifestyle behaviors likely to prevent excessive weight gain over time is particularly important in this college-aged population given that these behaviors will likely track into adulthood [1].

As noted, there were no significant differences between conditions in body weight at the end of the short intervention period; both groups were weight stable. However, students were at a healthy weight at baseline, and intervention materials were focused on maintaining a healthy weight; therefore, this is not completely unexpected. Furthermore, the intervention was implemented in the spring semester, a time when people are often making New Year's resolutions to improve their health behaviors [56]. It is unknown whether weight differences would have emerged if the sample had been followed for a longer period. Most of the limited number of weight gain prevention interventions that have been conducted targeting college students have been short-term interventions (6-15 weeks), and their findings are mixed. For example, Dennis et al [36] observed significant increases in weight among freshmen women engaged in two different SCT-based weight gain prevention interventions, which were delivered through the Internet as part of a course. Providing monetary incentives for remaining weight stable in conjunction with a behavioral intervention resulted in weight changes comparable to those resulting from the behavioral intervention only, without incentives. Average 14-week weight gains were 1.75 and 0.95 kg, respectively. Without a control group in this study, it is unclear whether the interventions attenuated weight gain; evidently, they did not prevent it. In contrast, Levitsky et al [35] included control groups against which to compare two different semester-long interventions characterized by daily weight monitoring with emailed feedback targeted for freshmen women. They found that the control groups gained significantly more weight in each of the two studies (3.1 and 2.0 kg) relative to the experimental groups (0.1 and –0.82 kg), which had negligible gains or modest losses over the 12-week follow-up. Weight gains among those women in the intervention groups were similar to those observed in this study, but the control group participants in the study by Levitsky et al [35] experienced markedly higher weight gain than those in this study. This may reflect the different sample characteristics between the two studies. Levitsky et al [35] enrolled freshmen women exclusively, whereas the sample population in this study was comprised almost entirely of upperclassmen and included both men and women. The rate of weight gain among college students can vary, and freshmen typically gain weight at a higher rate than upperclassmen [7,8,57], which could explain, in part, the difference in weight gains in the control groups observed between the study by Levitsky et al [35] and this study. Indeed, even differences in the timing of the data collection can result in different weight gain trajectories; the first semester of freshman year appears to have the greatest weight gains [9].

This study followed upperclassmen during the spring semester, which perhaps accounts for the minimal weight gain in the control group over the 9-week period. In one of the few longer-duration studies, an intervention focused on maintaining a healthy lifestyle that was delivered to first- and second-year college students in small-group seminars produced small weight losses over 2 years [30], whereas the control group gained weight, resulting in a significant net difference of 1.3 kg between the two groups. This suggests that with extended follow-up periods, the effect of weight gain prevention interventions may be more likely to surface. More definitive studies are required to determine when interventions need to be delivered for college students and how long they should last to achieve the best weight gain prevention outcomes.

Study Limitations and Strengths

Limitations of this study must be considered when interpreting the results. First, the sample size was small and relatively homogeneous, limiting the generalizability of the findings. The use of a self-report measure of weight control behaviors is another significant study limitation, as is the short intervention exposure and the limited follow-up period. The risk that contamination between conditions occurred must also be considered since students were enrolled at the same university and were aware of the content being covered in both the interventions as a result of reviewing the consent forms. Confounding due to contamination that resulted in an increased focus on weight management among controls could explain the lack of a weight change difference. However, the failure to find an increase in behavioral weight control strategies among controls suggests contamination across conditions was unlikely. Furthermore, the small number of classes randomized is another limitation of this pilot study and points to the need for a larger study of longer duration.

Nevertheless, there are several notable strengths of this pilot study that fuel enthusiasm for the intervention approach implemented. In what is the first report of a behavioral weight gain prevention program for college students that incorporated social media and mHealth monitoring devices, student engagement was robust and treatment satisfaction was high. Moreover, retention was excellent. The use of objective measures of weight and intervention engagement and assessment of both appropriate and inappropriate weight control behaviors represent additional study strengths. The automated delivery of the electronic newsletters and Facebook posts is an additional advantage that points to the potential to readily bring the intervention to scale should it prove efficacious in larger trials over a more extended period.

Conclusions

An online social media-based weight gain prevention intervention accompanied by mHealth self-monitoring tools increased appropriate weight control efforts by college students relative to controls and demonstrated no short-term iatrogenic effects, although weight change differences between groups were not apparent over the 9-week observation period. Experiences with the intervention indicate students readily engaged with all the technological platforms implemented and found the intervention acceptable. These preliminary findings

support the need for the evaluation of this type of intervention over a longer duration in order to determine whether engagement and adherence, as well as the observed behavioral improvements, can be sustained, and in turn, positively affect weight outcomes.

Acknowledgments

Authors' Contributions

DSW conceived the study, was the principal investigator, and took the lead in study design and implementation, as well as in drafting the manuscript. CMM contributed to the study design and was responsible for survey data collection design and implementation, as well as implementing social media elements of the healthy weight intervention, conducting analyses and drafting the discussion section of the manuscript. GT-M assisted in study design, provided content to the healthy weight intervention, and implemented the social media aspects of the intervention, as well as edited the manuscript. BS assisted with study design, was responsible (with HMB) for the design and implementation of the control intervention (HPV vaccination) materials, and edited the manuscript. CL developed and implemented delivery of the intervention newsletters, as well as social media elements of the healthy weight intervention, compiled engagement data, and edited the manuscript. KM was the study coordinator and participated in the overall design and implementation of the study, as well as edited the manuscript. SW provided study design contributions, including intervention design, and edited the manuscript. HMB provided content and oversight of the control intervention (HPV vaccination awareness), assisted in study design, and edited the manuscript. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Weight Control Practices Questionnaire.

[PDF File (Adobe PDF File), 72KB - jmir_v18i6e133_app1.pdf]

References

1. Nelson MC, Story M, Larson NI, Neumark-Sztainer D, Lytle LA. Emerging adulthood and college-aged youth: an overlooked age for weight-related behavior change. Obesity (Silver Spring) 2008 Oct;16(10):2205-2211. [doi: 10.1038/oby.2008.365] [Medline: 18719665]
2. Zamora D, Gordon-Larsen P, Jacobs DR, Popkin BM. Diet quality and weight gain among black and white young adults: the Coronary Artery Risk Development in Young Adults (CARDIA) Study (1985-2005). Am J Clin Nutr 2010 Oct;92(4):784-793 [FREE Full text] [doi: 10.3945/ajcn.2010.29161] [Medline: 20685947]
3. Yong LC, Kuller LH, Rutan G, Bunker C. Longitudinal study of blood pressure: changes and determinants from adolescence to middle age. The Dormont High School follow-up study, 1957-1963 to 1989-1990. Am J Epidemiol 1993 Dec 1;138(11):973-983. [Medline: 8256782]
4. Ford ES, Williamson DF, Liu S. Weight change and diabetes incidence: findings from a national cohort of US adults. Am J Epidemiol 1997 Aug 1;146(3):214-222 [FREE Full text] [Medline: 9247005]
5. Poirier P, Giles TD, Bray GA, Hong Y, Stern JS, Pi-Sunyer FX, et al. Obesity and cardiovascular disease: pathophysiology, evaluation, and effect of weight loss. Arterioscler Thromb Vasc Biol 2006 May;26(5):968-976 [FREE Full text] [doi: 10.1161/01.ATV.0000216787.85457.f3] [Medline: 16627822]
6. Vella-Zarb RA, Elgar FJ. The 'freshman 5': a meta-analysis of weight gain in the freshman year of college. J Am Coll Health 2009;58(2):161-166. [doi: 10.1080/07448480903221392] [Medline: 19892653]
7. Gropper SS, Simmons KP, Connell LJ, Ulrich PV. Changes in body weight, composition, and shape: a 4-year study of college students. Appl Physiol Nutr Metab 2012 Dec;37(6):1118-1123. [doi: 10.1139/h2012-139] [Medline: 22978391]
8. Gropper SS, Simmons KP, Connell LJ, Ulrich PV. Weight and Body Composition Changes during the First Three Years of College. J Obes 2012;2012:634048 [FREE Full text] [doi: 10.1155/2012/634048] [Medline: 23050127]
9. Lloyd-Richardson EE, Bailey S, Fava JL, Wing R, Tobacco Etiology Research Network (TERN). A prospective study of weight gain during the college freshman and sophomore years. Prev Med 2009 Mar;48(3):256-261 [FREE Full text] [doi: 10.1016/j.ypmed.2008.12.009] [Medline: 19146870]
10. Fedewa MV, Das BM, Evans EM, Dishman RK. Change in weight and adiposity in college students: a systematic review and meta-analysis. Am J Prev Med 2014 Nov;47(5):641-652. [doi: 10.1016/j.amepre.2014.07.035] [Medline: 25241201]

11.   American College Health Association. Undergraduate Students - Executive Summary Spring 2014. 2014. URL: http://www.acha-ncha.org/docs/NCHA-II_WEB-PAPER_SPRING2014_UNDERGRAD_REFERENCEGROUP_EXECUTIVESUMMARY.pdf [accessed 2015-11-19] [WebCite Cache ID 6d9uGHmPL]

12.   Wadden TA, Webb VL, Moran CH, Bailer BA. Lifestyle modification for obesity: new developments in diet, physical activity, and behavior therapy. Circulation 2012 Mar 6;125(9):1157-1170 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.111.039453] [Medline: 22392863]

13.   Smith A, Rainie L, Zickuhr K. College Students and Technology. 2011 Jul 19. URL: http://www.pewinternet.org/2011/07/19/college-students-and-technology/ [WebCite Cache ID 6d9wM18V0]

14.   Pearson. Pearson Student Mobile Device Survey 2014. 2014 May 16. URL: http://www.pearsoned.com/wp-content/uploads/Pearson-HE-Student-Mobile-Device-Survey-PUBLIC-Report-051614.pdf [accessed 2015-11-19] [WebCite Cache ID 6d9vhBEbf]

15.   Duggan M. The Demographics of Social Media Users. URL: http://www.pewinternet.org/2015/08/19/the-demographics-of-social-media-users/ [accessed 2016-05-05] [WebCite Cache ID 6hHcXcJcB]

16.   Cadmus-Bertram LA, Marcus BH, Patterson RE, Parker BA, Morey BL. Randomized Trial of a Fitbit-Based Physical Activity Intervention for Women. Am J Prev Med 2015 Sep;49(3):414-418. [doi: 10.1016/j.amepre.2015.01.020] [Medline: 26071863]

17.   Gilmore LA, Duhé AF, Frost EA, Redman LM. The technology boom: a new era in obesity management. J Diabetes Sci Technol 2014 May;8(3):596-608 [FREE Full text] [doi: 10.1177/1932296814525189] [Medline: 24876625]

18.   Steinberg DM, Tate DF, Bennett GG, Ennett S, Samuel-Hodge C, Ward DS. The efficacy of a daily self-weighing weight loss intervention using smart scales and e-mail. Obesity (Silver Spring) 2013 Sep;21(9):1789-1797 [FREE Full text] [doi: 10.1002/oby.20396] [Medline: 23512320]

19.   Ha E, Caine-Bish N. Effect of nutrition intervention using a general nutrition course for promoting fruit and vegetable consumption among college students. J Nutr Educ Behav 2009;41(2):103-109. [doi: 10.1016/j.jneb.2008.07.001] [Medline: 19304255]

20.   Hekler EB, Gardner CD, Robinson TN. Effects of a college course about food and society on students' eating behaviors. Am J Prev Med 2010 May;38(5):543-547. [doi: 10.1016/j.amepre.2010.01.026] [Medline: 20227847]

21.   Abu-Moghli FA, Khalaf IA, Barghoti FF. The influence of a health education programme on healthy lifestyles and practices among university students. Int J Nurs Pract 2010 Feb;16(1):35-42. [doi: 10.1111/j.1440-172X.2009.01801.x] [Medline: 20158546]

22.   Peterson S, Duncan DP, Null DB, Roth SL, Gill L. Positive changes in perceptions and selections of healthful foods by college students after a short-term point-of-selection intervention at a dining hall. J Am Coll Health 2010;58(5):425-431. [doi: 10.1080/07448480903540457] [Medline: 20304754]

23.   Brown KN, Wengreen HJ, Vitale TS, Anderson JB. Increased self-efficacy for vegetable preparation following an online, skill-based intervention and in-class tasting experience as a part of a general education college nutrition course. Am J Health Promot 2011;26(1):14-20. [doi: 10.4278/ajhp.091214-QUAN-389] [Medline: 21879938]

24.   Boyle J, Mattern CO, Lassiter JW, Ritzler JA. Peer 2 peer: efficacy of a course-based peer education intervention to increase physical activity among college students. J Am Coll Health 2011;59(6):519-529. [doi: 10.1080/07448481.2010.523854] [Medline: 21660807]

25.   Grim M, Hortz B, Petosa R. Impact evaluation of a pilot web-based intervention to increase physical activity. Am J Health Promot 2011;25(4):227-230. [doi: 10.4278/ajhp.081216-ARB-307] [Medline: 21361806]

26.   Tully MA, Cupples ME. UNISTEP (university students exercise and physical activity) study: a pilot study of the effects of accumulating 10,000 steps on health and fitness among university students. J Phys Act Health 2011 Jul;8(5):663-667. [Medline: 21734311]

27.   Magoc D, Tomaka J, Bridges-Arzaga A. Using the web to increase physical activity in college students. Am J Health Behav 2011;35(2):142-154. [Medline: 21204677]

28.   Martens MP, Buscemi J, Smith AE, Murphy JG. The short-term efficacy of a brief motivational intervention designed to increase physical activity among college students. J Phys Act Health 2012 May;9(4):525-532. [Medline: 21934167]

29.   Stice E, Orjada K, Tristan J. Trial of a psychoeducational eating disturbance intervention for college women: a replication and extension. Int J Eat Disord 2006 Apr;39(3):233-239. [doi: 10.1002/eat.20252] [Medline: 16498589]

30.   Hivert M, Langlois M, Bérard P, Cuerrier J, Carpentier AC. Prevention of weight gain in young adults through a seminar-based intervention program. Int J Obes (Lond) 2007 Aug;31(8):1262-1269. [doi: 10.1038/sj.ijo.0803572] [Medline: 17356531]

31.   Napolitano MA, Hayes S, Bennett GG, Ives AK, Foster GD. Using Facebook and text messaging to deliver a weight loss program to college students. Obesity (Silver Spring) 2013 Jan;21(1):25-31. [doi: 10.1002/oby.20232] [Medline: 23505165]

32.   Harvey-Berino J, Pope L, Gold BC, Leonard H, Belliveau C. Undergrad and overweight: an online behavioral weight management program for college students. J Nutr Educ Behav 2012;44(6):604-608. [doi: 10.1016/j.jneb.2012.04.016] [Medline: 23140565]

XSL•FO

RenderX

33. Davy BM, Potter KL, Dennis Parker EA, Harden S, Hill JL, Halliday TM, et al. Feasibility, effectiveness, and perceptions of an Internet-and incentive-based behavioral weight loss intervention for overweight and obese college freshmen: A mixed methods approach. OJPM 2013;03(07):429-440. [doi: 10.4236/ojpm.2013.37058]

34. Gow RW, Trace SE, Mazzeo SE. Preventing weight gain in first year college students: an online intervention to prevent the "freshman fifteen". Eat Behav 2010 Jan;11(1):33-39 [FREE Full text] [doi: 10.1016/j.eatbeh.2009.08.005] [Medline: 19962118]

35. Levitsky DA, Garay J, Nausbaum M, Neighbors L, Dellavalle DM. Monitoring weight daily blocks the freshman weight gain: a model for combating the epidemic of obesity. Int J Obes (Lond) 2006 Jun;30(6):1003-1010. [doi: 10.1038/sj.ijo.0803221] [Medline: 16446748]

36. Dennis EA, Potter KL, Estabrooks PA, Davy BM. Weight Gain Prevention for College Freshmen: Comparing Two Social Cognitive Theory-Based Interventions with and without Explicit Self-Regulation Training. J Obes 2012;2012:803769 [FREE Full text] [doi: 10.1155/2012/803769] [Medline: 22778919]

37. Laska MN, Pelletier JE, Larson NI, Story M. Interventions for weight gain prevention during the transition to young adulthood: a review of the literature. J Adolesc Health 2012 Apr;50(4):324-333 [FREE Full text] [doi: 10.1016/j.jadohealth.2012.01.016] [Medline: 22443834]

38. Bandura A. Health promotion from the perspective of social cognitive theory. Psychology & Health 1998 Jul;13(4):623-649. [doi: 10.1080/08870449808407422]

39. Diabetes Prevention Program (DPP) Research Group. The Diabetes Prevention Program (DPP): description of lifestyle intervention. Diabetes Care 2002 Dec;25(12):2165-2171 [FREE Full text] [Medline: 12453955]

40. Hales SB, Davidson C, Turner-McGrievy GM. Varying social media post types differentially impacts engagement in a behavioral weight loss intervention. Transl Behav Med 2014 Dec;4(4):355-362 [FREE Full text] [doi: 10.1007/s13142-014-0274-z] [Medline: 25584084]

41. 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 [FREE Full text] [doi: 10.1016/j.ypmed.2010.04.018] [Medline: 20478333]

42. Wing RR, Tate DF, Gorin AA, Raynor HA, Fava JL. A self-regulation program for maintenance of weight loss. N Engl J Med 2006 Oct 12;355(15):1563-1571. [doi: 10.1056/NEJMoa061883] [Medline: 17035649]

43. Linde JA, Jeffery RW, French SA, Pronk NP, Boyle RG. Self-weighing in weight gain prevention and weight loss trials. Ann Behav Med 2005 Dec;30(3):210-216. [doi: 10.1207/s15324796abm3003_5] [Medline: 16336072]

44. Vanwormer JJ, French SA, Pereira MA, Welsh EM. The impact of regular self-weighing on weight management: a systematic literature review. Int J Behav Nutr Phys Act 2008;5:54 [FREE Full text] [doi: 10.1186/1479-5868-5-54] [Medline: 18983667]

45. Wing RR, Tate D, LaRose JG, Gorin AA, Erickson K, Robichaud EF, et al. Frequent self-weighing as part of a constellation of healthy weight control practices in young adults. Obesity (Silver Spring) 2015 May;23(5):943-949. [doi: 10.1002/oby.21064] [Medline: 25865175]

46. Steinberg DM, Tate DF, Bennett GG, Ennett S, Samuel-Hodge C, Ward DS. Daily self-weighing and adverse psychological outcomes: a randomized controlled trial. Am J Prev Med 2014 Jan;46(1):24-29. [doi: 10.1016/j.amepre.2013.08.006] [Medline: 24355668]

47. 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]

48. Raynor HA, Jeffery RW, Ruggiero AM, Clark JM, Delahanty LM, Look AHEAD (Action for Health in Diabetes) Research Group. Weight loss strategies associated with BMI in overweight adults with type 2 diabetes at entry into the Look AHEAD (Action for Health in Diabetes) trial. Diabetes Care 2008 Jul;31(7):1299-1304 [FREE Full text] [doi: 10.2337/dc07-2295] [Medline: 18375417]

49. Look AHEAD Research Group. Eight-year weight losses with an intensive lifestyle intervention: the look AHEAD study. Obesity (Silver Spring) 2014 Jan;22(1):5-13 [FREE Full text] [doi: 10.1002/oby.20662] [Medline: 24307184]

50. Mackey E, Schweitzer A, Hurtado ME, Hathway J, DiPietro L, Lei KY, et al. The feasibility of an e-mail-delivered intervention to improve nutrition and physical activity behaviors in African American college students. J Am Coll Health 2015;63(2):109-117. [doi: 10.1080/07448481.2014.990971] [Medline: 25611932]

51. Rote AE, Klos LA, Brondino MJ, Harley AE, Swartz AM. The Efficacy of a Walking Intervention Using Social Media to Increase Physical Activity: A Randomized Trial. J Phys Act Health 2015 Jun;12 Suppl 1:S18-S25. [doi: 10.1123/jpah.2014-0279] [Medline: 25599378]

52. Matvienko O, Lewis DS, Schafer E. A college nutrition science course as an intervention to prevent weight gain in female college freshmen. J Nutr Educ 2001;33(2):95-101. [Medline: 12031189]

53. Cavallo DN, Tate DF, Ries AV, Brown JD, DeVellis RF, Ammerman AS. A social media-based physical activity intervention: a randomized controlled trial. Am J Prev Med 2012 Nov;43(5):527-532 [FREE Full text] [doi: 10.1016/j.amepre.2012.07.019] [Medline: 23079176]

54. Merchant G, Weibel N, Patrick K, Fowler JH, Norman GJ, Gupta A, et al. Click "like" to change your behavior: a mixed methods study of college students' exposure to and engagement with Facebook content designed for weight loss. J Med Internet Res 2014;16(6):e158 [FREE Full text] [doi: 10.2196/jmir.3267] [Medline: 24964294]

XSL•FO

RenderX

55.   Moreno JP, Johnston CA, Woehler D. Changes in weight over the school year and summer vacation: results of a 5-year longitudinal study. J Sch Health 2013 Jul;83(7):473-477. [doi: 10.1111/josh.12054] [Medline: 23782089]

56.   Norcross JC, Mrykalo MS, Blagys MD. Auld Lang Syne: Success Predictors, Change Processes, and Self-Reported Outcomes of New Year's Resolvers and Nonresolvers. Journal of Clinical Psychology 2002;58(4):397-405.

57.   Kelly MC, Latner JD. Evaluating patterns of weight and body composition change among college women. Eat Behav 2015 Apr;17:157-162. [doi: 10.1016/j.eatbeh.2015.01.016] [Medline: 25796066]

Abbreviations

BMI: body mass index
HW: healthy weight
HPV: human papillomavirus
SCT: social cognitive theory

XSL•FO

RenderX

Original Paper

Changing Behavioral Lifestyle Risk Factors Related to Cognitive Decline in Later Life Using a Self-Motivated eHealth Intervention in Dutch Adults

Teun Aalbers[1,2], PhD; Li Qin[1,2], PhD; Maria AE Baars[3], PhD; Annet de Lange[4,5], PhD; Roy PC Kessels[6,7], PhD; Marcel GM Olde Rikkert[1,2], PhD

[1]Radboud University Medical Center, Department of Geriatric Medicine, Nijmegen, Netherlands

[2]Radboud University Medical Center, Radboud Alzheimer Center, Nijmegen, Netherlands

[3]HAN University of Applied Sciences, Institute of Social Sciences, Nijmegen, Netherlands

[4]HAN University of Applied Sciences, Faculty of Human Resource Management, Nijmegen, Netherlands

[5]Norwegian School of Hotel Management, University of Stavanger, Stavanger, Norway

[6]Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, Netherlands

[7]Radboud University Medical Center, Department of Medical Psychology, Nijmegen, Netherlands

Corresponding Author:
Teun Aalbers, PhD
Radboud University Medical Center
Department of Geriatric Medicine
Reinier Postlaan 4
Nijmegen, 6500 HB
Netherlands
Phone: 31 243619807
Fax: 31 243617408
Email: teun.aalbers@radboudumc.nl

Abstract

Background: Our labor force is aging, but aged workers are not yet coached on how to stay cognitively fit for the job.

Objective: In this study, we tested whether a self-motivated, complex eHealth intervention could improve multiple health-related behaviors that are associated with cognitive aging among working Dutch adults.

Methods: This quasi-experimental prospective study with a pre-post design was conducted with employees of Dutch medium to large companies. All employees with Internet access, a good understanding of the Dutch language, and who provided digital informed consent were eligible to participate. In total, 2972 participants (2110/2972, 71.11% females) with a mean (standard deviation, SD) age of 51.8 (SD 12.9) years were recruited; 2305 became active users of the intervention, and 173 completed the 1-year follow-up. This self-motivated eHealth lifestyle intervention stimulates participants to set personally relevant, monthly health behavior change goals using Goal Attainment Scaling and to realize these goals by implementing behavior change techniques grounded in behavior change theory. The primary outcomes were the goal-setting success rate and the change in overall lifestyle score from baseline to the 1-year follow-up; the score was based on physical activity, diet, smoking, alcohol, sleep, and stress scores. The secondary outcomes were the changes in body weight, body mass index, specific lifestyle characteristics, and website usage.

Results: A total of 1212 participants set 2620 behavior change goals; 392 participants assessed 1089 (1089/2288, 47.59%) goals and successfully achieved 422 (422/1089, 38.75%) of these goals. Among the goal-setting participants in follow-up, this led to a +0.81-point improvement (95% CI 0.49-1.13, $P<.001$) in overall lifestyle (d=0.32) and weight loss of 0.62 kg (95% CI −1.16 to −0.07, P=.03). These participants also showed significant improvement in 8 out of 11 specific lifestyle components.

Conclusions: Among an adult Dutch population, this eHealth intervention resulted in lifestyle changes in behavioral risk factors associated with cognitive decline, and these improvements lasted over the period of 1 year. Given the general aging of our workforce, this eHealth intervention opens new avenues for the widespread use of cost-effective self-motivated prevention programs aimed at prevention of early-stage cognitive decline and more self-management of their risk factors.

XSL•FO
RenderX

Introduction

A number of large-scale longitudinal studies have shown that several behavioral risk factors are associated with the onset and progression of diabetes, cardiovascular disease, stroke, and cognitive impairment [1-3]. Moreover, studies have shown that the relative risk of developing these diseases increases when several risk factors are present [4]. Consequently, reversing unhealthy behaviors may significantly benefit one's health in later life. Three key unhealthy behaviors—a lack of physical activity, consuming an unhealthy diet, and tobacco use—are estimated to account for approximately 71% of the more than 1 million preventable deaths that occurred in the year 2000 in the United States alone [5,6]. Moreover, a growing body of evidence suggests that good health during one's midlife years has a positive effect on cognitive aging, and focusing on modifiable lifestyle-related risk factors can delay or even prevent the onset of Alzheimer disease [7-9]. Physical activity, nutrition, adiposity, smoking, alcohol consumption, sleep, and stress have all been identified as modifiable lifestyle factors that are associated with cognitive aging [10-18]. A recent study even predicted that one-third of all global cases of Alzheimer disease can be attributed to potentially modifiable risk factors [19], and according to a study by Barnes and Yaffe [7], even a 10%-25% reduction in modifiable risk factors for Alzheimer disease might prevent up to 3 million cases of Alzheimer disease worldwide.

Internet-Based Lifestyle Programs

In a recent systematic review, we found that tailored Internet-based intervention programs provide an evidence-based means to effect large-scale change in modifiable risk factors [20]. The efficacy and feasibility of Web-delivered intervention programs for changing unhealthy behaviors is well established with respect to the relevant health behaviors (eg, increased physical activity, weight loss, smoking cessation, and reduced alcohol consumption) [21,22]. Moreover, more recent research attempted to determine which types of interventions are the most effective. For example, recent meta-analyses revealed that some form of tailoring is necessary for improving the personal relevance and extent of health behavior change programs [23] and for increasing the intervention's effectiveness [24]. However, to date, no published eHealth study has been designed to investigate the effect of lifestyle changes on cognitive aging.

Objective of the Study

Our objective was to design an innovative eHealth intervention program that motivates aging adults to adopt healthy lifestyle changes in order to help prevent cognitive decline. To facilitate feedback and to increase participant motivation, the intervention enables participants to monitor their own cognitive functioning over time by playing applied games (abbreviated here as the BAM-COG, Brain Aging Monitor–Cognitive Assessment Battery, part of the intervention), which provide a valid measure of various cognitive functions [25]. These games are part of the Brain Aging Monitor (BAM), an eHealth system with minimal barriers for participation, and they are sustainable in a wide range of practice settings. After developing and pilot-testing this eHealth intervention program, we initiated this study by asking whether using this self-motivated, complex eHealth intervention could effectively result in changing multiple health behaviors at midlife that are associated with cognitive decline in later life.

Methods

Study Design

This study used a pre-post design, in which newly enrolled participants chose their own time path, from registration to setting goals and monitoring their change in behavior. Randomizing the participants against a sham intervention in an occupational health promotion program, in which the stimulated behaviors are known to be advantageous, was judged to be insufficiently motivating for the participants and potential participating companies; therefore, we chose to use a pre-post design. Inclusion started in October 2012. This intervention was registered with the Dutch Trial Register (NTR4144) and was exempt from formal testing by the Medical Ethics Committee of the Radboud university medical center Nijmegen, which determined that the intervention was not invasive, risky, or burdensome.

Study Population

Participants were recruited from medium to large companies and from the general population. Due to the Internet-based nature of the eHealth intervention, no regional restrictions applied. Although the intervention was primarily aimed at participants aged 40 years and older, no age restriction for registration applied, as this did not result in any additional logistics or cost. Participants were required to have regular Internet access at their work and/or home. This was not a relevant barrier to participation, as approximately 92% of individuals aged 45-75 years in the Netherlands have Internet access [26]. Because the intervention was available only in Dutch, a good understanding of the Dutch language was a prerequisite for registration, and all participants were required to provide electronic written informed consent. Because the presented outcomes are intermediate outcomes for the overall 2-year follow-up data outcomes, no power analysis was performed specifically for the reported outcomes.

XSL•FO
RenderX

Assessment of Risk Factors Related to Cognitive Aging

At baseline, physical activity, nutrition, smoking, alcohol consumption, sleep patterns, and stress behavior data were collected using electronic questionnaires. For a detailed overview of the room for improvement in these 6 lifestyle factors in Dutch society, see Multimedia Appendix 1: Room for improvement in the Netherlands. We monitored the participants' cognitive functions (eg, working memory, visuospatial short-term memory, and planning performance) using the BAM-COG, an Internet-based tool for self-monitoring cognitive functioning using applied games that we previously developed and validated [25]. In addition, we administered the Dutch General Self-Efficacy Scale [27], lifestyle factor–specific self-efficacy questions [28], the Positive and Negative Affect Schedule [29], and the Self-Control Scale [30]. A complete overview of these questionnaires is available from the protocol [31]. Twelve months after starting the intervention, the participants were automatically prompted by email to repeat the e-questionnaires and BAM-COG. We monitored the number of log-in events, the number of goals set, the number of goals assessed (ie, the goals that were scored by the participant), and whether or not a goal was achieved. Goals were set using Goal Attainment Scaling (GAS; see the example in Table 1; number of days and length of exercise bouts are variable) [32]. The main features and strengths of using GAS are that it enables BAM to compare goals over different lifestyle modalities, it provides positive feedback regarding partially accomplished goals, and it stimulates the participant to consciously consider what goals are realistic within a given time frame.

Table 1. One possible example of a filled out Goal Attainment Scaling for the goal "I want to exercise more."

	Behavior frequency		Behavior duration	GAS[a] score
I fall short of my goal if I exercise	1 time per week	For	45 minutes	−2
I fall a little short of my goal if I exercise	2 times per week	For	45 minutes	−1
I reach my goal if I exercise	3 times per week	For	45 minutes	0
I exceed my goal if I exercise	4 times per week	For	45 minutes	+1
I greatly exceed my goal if I exercise	5 times per week	For	45 minutes	+2

[a] GAS: Goal Attainment Scaling.

Using these GAS scores, we measured both overall success and the success of each specific lifestyle area (as the percentage of achieved goals out of the total number of goals set). To measure overall lifestyle change, we calculated an overall lifestyle score based on the following 8 lifestyle measures: physical activity, exercise, healthy nutritional behavior, unhealthy nutritional behavior, smoking status, alcohol consumption, sleep status, and stress status. Each of these factors was categorical with a value of 1-3, thus summing to a total score that ranged from 8 (ie, an unhealthy lifestyle) to 24 (ie, a healthy lifestyle). More detailed information regarding how the values were defined for each lifestyle factor is provided in Multimedia Appendix 2: Construction of the overall lifestyle score. It is important to note that an outcome on a GAS score does not directly match a change in category in Multimedia Appendix 2. They are, however, related in such a way that if the participant in the example of Table 1 would obtain a GAS score of 0 or higher, he or she would switch from the category of "suboptimally active" to "norm active."

The Brain Aging Monitor Intervention

Participants were able to register free of charge at the intervention website. After providing electronic written informed consent and receiving email validation of their account, the participant could log on to a personalized "dashboard." After completing the questionnaires (for their tailored intervention content), each participant received a personalized lifestyle overview indicating room for improvement, after which the participant was invited to complete the 3 validated Internet-based puzzle games to assess their baseline cognitive performance.

After the questionnaires were completed, the intervention components were unlocked, thus enabling the participant to begin setting behavioral goals using the GAS methodology [32] in the BAM interface. After a participant set a goal, positive reinforcement was provided, along with practical tips and tricks to accomplish that specific goal (eg, when a participant's goal was to "start exercising," the tips and tricks provided included training schedules building up to a 5-km run or 500-m swim). Because using behavior change techniques that are based on evidence-based principles leads to better final scores [33], we incorporated 13 of the 26 behavior change techniques that were identified in the taxonomy by Abraham and Michie [34]. These techniques are grounded in the Social Cognitive Theory [35,36], the Transtheoretical Model [37], the Theory of Reasoned Action, and the Theory of Planned Behavior [38]. Each goal was then transferred to the short-term monitoring system, in which the participants could monitor their behavior on a daily basis; behavior was represented graphically in bar charts. After 1 month, the participants were asked whether the goal was achieved. If the participant answered this question (regardless of the answer), the goal was considered assessed. If the result was positive (ie, a GAS score of ≥0), the goal was registered as being achieved successfully and the goal was transferred to the long-term monitoring system (LTMS); if the result was negative (ie, a GAS score of <0), the goal was deleted from the participant's profile. After a goal with a positive outcome was entered into the LTMS, it was monitored monthly in order to track the behavior and stimulate its maintenance.

The BAM automatically sent reminder emails to the participants each week. The frequency of these reminder emails could be changed by the participant to daily, biweekly, or monthly

XSL•FO

RenderX

intervals. In addition to the goal-setting and monitoring components, the BAM also featured weekly blogs and healthy recipes [31].

Primary and Secondary Outcomes

We defined changes in risk factors for cognitive aging and website use as the 1-year outcomes. The cognitive outcome measures require longer follow-up (eg, 2 years) and were not available at the time of publication [31]. The first 1-year primary outcome was the overall and lifestyle-specific goal-setting success rates, which were calculated as GAS scores of ≥0. The second primary outcome was the change in overall lifestyle score from baseline to the 1-year follow-up. The secondary outcomes were changes in body weight, body mass index (BMI), and the resulting changes in specific lifestyle areas.

Data Analyses

The differences between the group of participants who set goals after doing the pretest (goal-setting group) and the group of participants who did not set goals after the pretest (non–goal-setting group) at baseline were analyzed using the independent samples *t* test (for interval variables) or the χ^2 test (for categorical variables). The Mann-Whitney *U* test was used to compare differences in intervention usage. We used the paired samples *t* test to analyze within-group differences in overall lifestyle score, BMI, weight, and lifestyle-specific areas. An analysis of covariance (ANCOVA) was used to calculate the mean change in body weight, BMI, and lifestyle changes, as

well as the 95% CI. For changes in body weight and BMI, the baseline values were adjusted as covariates in the ANCOVA in order to control for a potential regression-to-the-mean effect. After adjusting for covariates, multivariate linear regression analyses were performed to assess the linear associations between the total number of goals set (as a proxy for intervention utilization) and the changes in lifestyle factors at the 1-year follow-up. Unless stated otherwise, the outcome values are presented as the mean and standard deviation (SD); where possible, 95% CI is presented as well. Effect size (Cohen's *d*) was calculated for the overall lifestyle change. All *P* values are based on 2-sided testing, and all statistical analyses were performed using SPSS version 20 (SPSS Inc, Chicago, IL, USA).

Results

Baseline Characteristics

A total of 2972 people registered via the website, of whom 2305 became active users (see the flowchart in Figure 1). The mean (SD) age at registration was 51.8 (SD 12.9) years, and 71% of the participants were female. After the baseline measurement, 1212 participants proceeded with setting behavior change goals. Thus, 1093 participants never set a behavior change goal. The participants who set goals were more likely to be female, were less likely to have completed secondary school, and reported less healthy nutrition, and their overall lifestyle score was lower (Table 2).

Figure 1. Flowchart of Brain Aging Monitor participants.
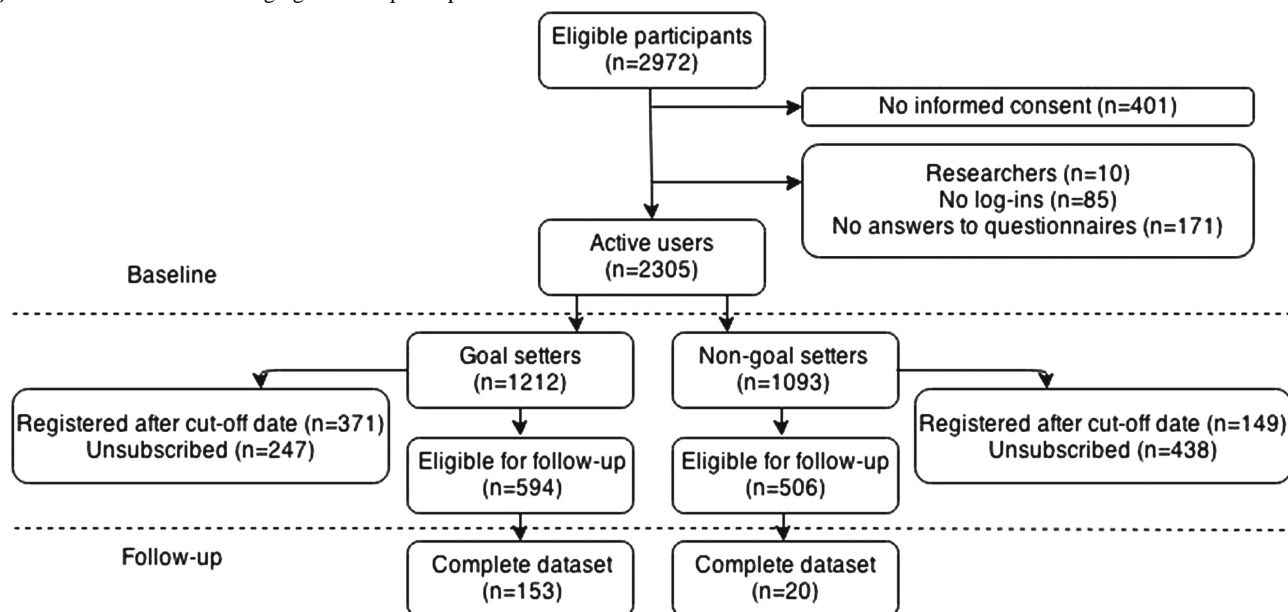
Table 2. Baseline characteristics of the participants who set goals and those who did not.

Characteristic[a]	Goal-setting group (n=1212)	Non–goal-setting group (n=1093)
Age in years, mean (SD)	52.34 (12.21)	51.28 (13.73)
Gender, female, n (%)	862 (71.11)	689 (63.04)[b]
Education level, n (%)		
Secondary school or lower	391 (32.3)	397 (37.2)[c]
Vocational degree	537 (44.3)	421 (39.5)
University degree	284 (23.4)	249 (23.3)
Body weight (kg), mean (SD)	75.5 (14.4)	75.6 (14.8)
BMI[d] (kg/m^2), mean (SD)	25.2 (4.2)	25.0 (4.4)
Subjective health, n (%)		
Poor	13 (1.1)	16 (1.6)
Fair	164 (13.5)	144 (13.5)
Good	796 (65.7)	654 (61.5)
Very good	187 (15.4)	188 (17.7)
Excellent	52 (4.3)	61 (5.7)
Physical activity, n (%)		
Inactive	274 (22.6)	206 (19.7)
Suboptimally active (1-4 days per week)	340 (28.1)	292 (27.9)
Norm active (≥5 days per week)	598 (49.3)	547 (52.3)
Exercise, n (%)		
Inactive	303 (25.0)	239 (22.9)
Suboptimally active (1 day per week)	235 (19.4)	210 (20.1)
Norm active (≥2 days per week)	674 (55.6)	596 (57.0)
Healthy nutrition, range 0-30, mean (SD)	23.6 (4.6)	23.6 (4.8)
Unhealthy nutrition, range 0-14, mean (SD)	5.8 (3.9)	5.4 (3.8)[e]
Smoking, n (%)		
Smoker	85 (7.0)	80 (7.7)
Ex-smoker	535 (44.1)	455 (44.0)
Nonsmoker	592 (48.8)	499 (48.3)
Alcohol, n (%)		
Abstainer	257 (21.2)	223 (21.6)
Drinker (1-5 days per week)	702 (57.9)	571 (55.3)
Frequent drinker (≥6 days per week)	253 (20.9)	238 (23.1)
Sleep pattern, n (%)		
Poor sleeper	268 (22.2)	208 (20.5)
Suboptimal sleeper	557 (46.1)	437 (43.1)
Good sleeper	384 (31.8)	369 (36.4)
Sleep hygiene score, mean (SD)	10.6 (4.0)	10.0 (4.2)
Satisfaction with life, range 5-35, mean (SD)	25.0 (6.3)	25.2 (6.3)
Overall lifestyle score, range 8-24, mean (SD)	17.4 (2.7)	17.7 (2.6)[c]
Personality questionnaires, mean (SD)		
DGSES[f] (range 10-40)	32.0 (4.3)	32.3 (4.8)

Characteristic[a]	Goal-setting group (n=1212)	Non–goal-setting group (n=1093)
Positive Affect (range 10-50)	33.9 (5.2)	33.7 (5.5)
Negative Affect (range 10-50)	21.1 (6.1)	20.4 (5.8)
Self-Control Scale (range 36-180)	125.6 (13.2)	126.7 (13.1)

[a] Mean (SD) is presented for continuous and ordinal variables, and n (%) is presented for categorical variables. For healthy nutrition, satisfaction with life, overall lifestyle score, DGSES, Positive Affect, and the Self-Control Scale, a higher score represents better performance; for unhealthy nutrition and Negative Affect, a lower score represents better performance.

[b] $P \leq .001$.

[c] $P < .05$.

[d] BMI: body mass index.

[e] $P < .01$.

[f] DGSES: Dutch General Self-Efficacy Scale.

Use of the Brain Aging Monitor for Goal Setting

The 2305 participants logged on to the BAM a total of 14,225 times. The non–goal-setting group logged on with a mean (SD) of 2.6 (SD 1.9) times per participant, which was significantly fewer than the goal-setting group (9.4, SD 21.3, visits per participant), even when 12 goal-setting participants who logged on >100 times each were excluded (resulting in 7.7, SD 10.0, visits per participant in the goal-setting group; $P < .001$, $r = .46$).

The 1212 goal-setting participants set a total of 2620 lifestyle goals (with 2.2, SD 3.6, goals set per participant). Of these 2620 goals, 2288 were predefined lifestyle goals in the GAS method, and the remaining 332 goals were created without the use of GAS. Of the 2288 prespecified behavior change goals, 1089 (1089/2288, 47.59%) were assessed, and 422 of the completed goals (422/1089, 38.75%) were achieved. Interestingly, all 1089 completed goals (goals that participants set and reported its status on) were completed by a subgroup of 392 participants (392/1212, 32.3% of the goal-setting group); the remaining 820 (820/1212, 67.7%) participants did not complete their goals (so did not report their status) within the study period. Table 3 summarizes the number of goals that the participants set, categorized by lifestyle area; Table 3 also presents statistics for

the goal-setting participants who reached their follow-up period and those who did not reach the follow-up period. An interesting result from Table 3 is the notion that goal-setting participants in follow-up had a very high completion rate (575/673, 85.4%) of goals, compared with the goal-setting participants who are not in follow-up (514/1615, 31.8%), but the goal achievement rate was roughly equal: 39.5% (227/575) versus 37.9% (195/514), respectively. Presumably, this is the case because participants who drop out no longer complete their goals and participants who reached their initial goals dropped out of the study, as they no longer perceived added value of the program.

On the reference date for follow-up (March 9, 2014), 1785 participants had been registered for at least 1 year. However, 685 of these participants withdrew from the program, leaving 1100 participants who could provide follow-up data (see Figure 1). In comparison with the 1059 goal-setting participants who were lost to follow-up, the 153 goal-setting participants who reached the follow-up phase had lower BMI ($P = .03$), were more educated ($P = .04$), were more likely to exercise ($P = .02$), and were more likely to be nonsmokers ($P = .04$) at baseline. However, together these differences did not result in a significant difference in overall lifestyle scores between these 2 groups ($P = .50$).

Table 3. Overview of goal setting, goal assessment, and goal achievement among the 1212 participants in the goal-setting group.

	Goals set[a], n	Goals completed[b], n (%)	Goals achieved[c], n (%)
Lifestyle area [d]			
Physical activity	590	270 (45.8)	131 (48.5)
Weight gain	7	1 (14.3)	0 (0.0)
Weight loss	516	243 (47.1)	57 (23.5)
Healthy nutrition	263	126 (47.9)	69 (54.8)
Unhealthy nutrition	288	143 (49.7)	65 (45.5)
Smoking	21	6 (28.6)	2 (33.3)
Alcohol consumption	184	114 (61.2)	47 (41.2)
Sleep	304	154 (50.7)	39 (25.3)
Stress	115	32 (27.8)	12 (37.5)
Total number of goals	2288	1089 (47.59)	422 (38.75)
Participants in follow-up [e]			
Total number of goals	673	575 (85.4)	227 (39.5)
Participants not in follow-up [f]			
Total number of goals	1615	514 (31.82)	195 (37.9)

[a] The number of goals set for each specific lifestyle area.

[b] Percentage of completed goals out of the number of goals set.

[c] Percentage of achieved goals out of the number of goals completed.

[d] The total number of participants in the group is 1212, and the number of participants who completed goals is 392.

[e] The number of participants who reached follow-up is 153 and, of those, the number of participants who completed goals is 127.

[f] The number of participants who did not reach follow-up is 1059 and, of those, the number of participants who completed goals is 265.

Overall and Lifestyle-Specific Lifestyle Changes

After 1 year of BAM intervention, the participants in the goal-setting follow-up group had a significant improvement in their overall lifestyle scores (with a mean change of +0.81, SD 1.92, 95% CI 0.49-1.13; $P<.001$, $d=0.32$). The overall improvement was even greater when only the participants who achieved their goals were taken into consideration (with a mean change of +1.01, SD 1.88, 95% CI 0.61-1.41; $P<.001$, $d=0.39$). Thus, remaining in the program and successfully reaching one's goals translates to a higher overall lifestyle score over 1 year's time. Table 4 summarizes the changes in lifestyle factors in the goal-setting group after 1 year in the program. We found that body weight, BMI, physical activity, healthy nutritional habits, unhealthy nutritional habits, hours slept per 24-hour period, and sleep hygiene were positively affected after 1 year's participation in the BAM. With respect to smoking, no results can be reported, as none of the users in the follow-up group completed smoking-related behavior change goals. With the exceptions of the sleep outcomes and satisfaction with life, the participants who set a goal had significantly lower baseline values in each specific lifestyle area than the participants who did not set goals in that specific lifestyle area. For example, at baseline, the participants who set an exercise goal averaged 1.5 days of exercise per week, whereas the participants who did not set an exercise goal were already exercising an average of 2.3 days per week ($P=.001$).

In addition, using multivariate linear regression analyses, we investigated the association between the total number of goals set by each participant and the change in lifestyle factors from baseline to the 1-year time point; this association was measured for the 173 participants who completed the 1-year follow-up. Our analysis revealed that setting more goals was significantly associated with the participants' ability to achieve weight loss (adjusted for gender and age at baseline; beta=−.08, 95% CI −.14 to −.02, $P=.01$), reduce BMI (adjusted for gender and age at baseline; beta=−.03, 95% CI −.05 to −.01, $P=.01$), and reduce unhealthy nutritional behavior (beta=−.06, 95% CI −.10 to −.02, $P=.01$). Setting more goals was also positively associated with an improvement in overall lifestyle (beta=.03, 95% CI −.001 to .07, $P=.06$) and healthy nutritional behavior (beta=.05, 95% CI −.001 to .01, $P=.06$). Finally, the total number of goals set was also positively correlated—albeit not significantly—with exercise, alcohol consumption, sleep hygiene, and sleep pattern (data not shown); in contrast, no such trend was found with respect to physical activity or satisfaction with life.

Table 4. Change in the lifestyle factors within the goal-setting group (n=153).

Cognitive aging risk factor	Baseline value mean (SD)	Value at year 1 mean (SD)	Mean change (95% CI)[a]	P value
Weight (kg)	74.9 (13.5)	74.3 (13.4)	−0.62 (−1.16 to −0.07)	.03
BMI[b] (kg/m^2)	24.6 (3.9)	24.4 (3.8)	−0.20 (−0.38 to −0.03)	.03
Physical activity (p/w[c])	4.1 (2.0)	4.3 (2.0)	+0.27 (0.02-0.52)	.04
Exercise (p/w)	1.9 (1.4)	2.1 (1.4)	+0.17 (−0.02 to 0.37)	.08
Healthy nutrition score	23.7 (3.9)	24.8 (4.0)	+1.09 (0.63-1.56)	<.001
Unhealthy nutrition score	5.7 (3.7)	5.0 (3.7)	−0.67 (−1.05 to −0.29)	.001
Alcohol consumption (p/w)	2.9 (2.4)	2.7 (2.3)	−0.13 (−0.31 to 0.04)	.14
Sleep (hours/24-hour period)	6.8 (0.9)	6.9 (0.9)	+0.12 (0.01-0.22)	.03
Sleep hygiene (points)	10.3 (4.1)	9.0 (4.5)	−1.32 (−1.85 to −0.78)	<.001
Satisfaction with life	24.5 (6.7)	24.2 (7.2)	−0.16 (−1.86 to 0.86)	.75
Overall lifestyle score	17.6 (2.6)	18.4 (2.6)	+0.81 (0.50-1.12)	<.001

[a] Each mean change and 95% CI was adjusted for gender and age at baseline, and weight and BMI were additionally adjusted for baseline weight.

[b] BMI: body mass index.

[c] p/w: per week.

Discussion

Here, we report that there is a plausible association between utilizing a self-motivated eHealth intervention program for 1 year and one's overall lifestyle scores by introducing lifestyle-specific health behavior changes that are relevant to cognitive decline. Although the effect size on overall lifestyle score at the individual participant level can be considered moderate (d=0.32-0.39) for a therapeutic intervention, from a public health perspective this effect size can be considered highly relevant and may deliver substantial added value to society [33].

Strengths of the Study

The primary strength of this study is that the intervention can reach and select the participants who score suboptimally in specific lifestyle areas, thereby motivating these specific individuals to set—and reach—realistic lifestyle goals and facilitating long-term health-related changes in behavior. This study therefore demonstrates proof of concept for the BAM intervention, focusing on improving cognitive aging risk factors among employees. Similar studies of other risk factors reported similar effect sizes [39] and concluded that these effects are common among computerized interventions; moreover, evidence suggests that these small to medium effect sizes can translate to large public health gains when implemented on a wide scale [23,24]. Therefore, the need for scalable lifestyle eHealth programs is clear, particularly given that smaller programs—although potentially more effective—lack the cost-effectiveness needed for large-scale public health implementation [40]. This improvement in public health is necessary, because epidemiological studies have found that healthy living—characterized by adherence to multiple healthy behavioral modalities—promotes positive physical and cognitive aging, whereas unhealthy living has a clear negative effect on both [1-3]. Because the BAM is targeted at the working

population (a group that is intrinsically self-motivated to stabilize and/or improve their working capacity, particularly in times of economic crisis) it may benefit large populations during this important age window. From a scalability point of view, the funds needed to implement, recruit, and administer the BAM intervention likely favor eHealth over more traditional, expert-led face-to-face interventions [41].

The multimodal nature of cognitive decline justifies choosing a multimodal intervention over single modality programs. Multiple behavior change programs address a broader scope of risk factors, delivering tailored and more comprehensive help to the participant. Moreover, the current public health status in the Netherlands (see Multimedia Appendix 1: Room for improvement in the Netherlands) emphasizes the added value provided by multiple health behavior change interventions. For this reason, we analyzed the effect of the intervention on separate lifestyle factors, as well as the overall lifestyle score. Because one's overall lifestyle is associated with cognitive function [9,12], measuring changes in overall lifestyle using an aggregated measure seems to make intuitive sense. In support of this approach, other studies measured overall lifestyle in order to create risk profiles for overall mortality, cardiovascular disease, cancer, and diabetes [4,42]. Compared with the relatively small average effect sizes reported by systematic reviews and meta-analyses of other tailored eHealth interventions that focus on more traditional health messages, d=0.19 [23], d=0.17 [24], and d=0.16 [33], our effect size with respect to overall lifestyle change was large (d=0.32-0.39). Our participants' weight loss was similar to previously reported interactive computer-based interventions, particularly given that most weight-loss interventions are aimed at overweight individuals instead of the general population that was targeted in this study [43].

The effectiveness may be explained—at least in part—by the fact that the BAM uses a novel motivation for healthy living

XSL•FO
RenderX

(eg, self-monitoring of cognition with games), which by itself has been advocated for many years [44]. However, the general public is generally unfamiliar with the notion that healthy living can affect cognitive health in later life; indeed, to the best of our knowledge, the BAM is the first eHealth intervention that is aimed in this direction. Our results suggest a trend toward a dose-response effect between the number of goals set and lifestyle changes. This is consistent with other studies in which utilization of the intervention predicts outcome measures [21,39]. Consistent with our results, a synthesis of meta-analyses and reviews found that single health behavior change programs were more effective at changing physical activity and dietary behavior, whereas multiple behavior change programs were more effective at inducing weight loss [45]. Given the results of our study, alternative routes to successfully change participants' physical activity and dietary behavior can be integrated into future eHealth interventions. Thus, although a program may initially be broad in its overall scope, it can subsequently focus on tailored behaviors. A systematic review by Nigg and Long [46] revealed a lack of multiple health behavior change interventions in older adults (ie, aged more than 55 years) for comparison with the effectiveness of single health behavior interventions, underscoring the need for interventions similar to BAM, so that further effectiveness comparisons can be done.

Limitations of the Study

This study also has some limitations. First, high dropout rates are a well-known limitation in eHealth research in general [21,39,47,48]. Although study retention does not necessarily affect the study's outcomes [23], low study adherence can hamper external validity and makes results prone to self-selection bias. In this case, the description of this selection bias is also a relevant study result as we meant to identify what subset of the general population would subscribe and adhere to the program and which results could be acquired in that population. There is a form of self-selection bias (eg, higher percentage of female participants) caused by the fact that it is a self-management tool. The BAM could not be forced on participants, but the study outcomes still may point at effectiveness for the type of participants that the intervention program will likely reach in later phases of implementation (ie, external validity). Therefore, it makes the per-protocol results relevant to people who are interested in self-management using eHealth, still resembling the Dutch population of the same age distribution. This in our view increases the external validity. In our study, the average number of times each participant logged on to the BAM is higher than in a similar study [39]; however, the BAM should be improved further in order to increase adherence, thereby optimizing the public health impact. Initial tailoring, repeat notifications, and monitoring are important factors for increasing adherence; however, maintaining the participants' interest is difficult with eHealth interventions that run longer than a few weeks [21]. Making the program more interactive and increasing the overall attractiveness of the design are two logical steps toward achieving higher adherence rates and increasing societal impact [49,50].

Second, our choice of a quasi-experimental pre-post design was driven by the predicted dropout rate and the setting of the study;

this design was the most feasible option for this type of pragmatic field trial [40]. Although a cluster randomized trial would have been preferred, such a design was not feasible given the low number of participating companies. For practical reasons, the companies agreed to include the BAM in their health care policy only if all of their employees would be allocated to the experimental group. Randomizing the participants at each site into control and treatment groups would have posed many practical problems (eg, blinding and allocation criteria [51]). Therefore, given the dropout rate, in an attempt to maximize statistical power by keeping the study group as large as possible, the use of a quasi-experimental design when performing longitudinal self-motivated eHealth research is justified [40]. For the same reason no a priori power analysis and sample size calculation was performed, because it would have been a poorly educated guess at best as we were unaware of the attrition for this innovative Internet-based approach, and because we would need several separate power analyses for the different lifestyle factors addressed. As this was not a randomized controlled trial, we could not follow the CONSORT (Consolidated Standards of Reporting Trials) guidelines at all points to lead us to the feasibility results we were looking for. However, we acknowledge the need for more controlled trials in order to further investigate the causal and dose-response relationship between the intervention uptake and lifestyle change. These trials will all need to find a way to optimally decrease the possible reporting bias that self-reported measures inherently bring in to the study design.

Third, these results are based on self-reported measures, which may have introduced a social desirability bias. It is known that self-reported height, weight, and BMI are reliable using eHealth [52]. The self-management nature and lack of researcher-participant interaction caused multiple participants to state that they "did not feel part of a research project," decreasing the likelihood of giving socially acceptable answers for the sake of the research team. However, no formal steps were taken to prevent participants from giving socially desirable answers. Moreover, self-reported measures provide a suitable reflection of the way in which participants view their own behavior and therefore serve as a suitable starting point for measuring behavioral changes that are perceived as personally relevant. Our recruitment strategies intentionally favored higher educated, "white collar" participants, as these participants would benefit most from a program with cognitive outcome measurements; such participants are likely to notice a small decline in executive functioning performance at work at a relatively young age. It may also be argued that the overrepresentation of women (71.1% of participants were female) may have affected the outcome of our study. However, this overrepresentation is expected in health-related Internet-based research studies, in which the average female participation rate was 64% [23]; moreover, neither age nor gender was significantly associated with the intervention outcome.

Finally, one could argue that the construction of the overall lifestyle score was—at least to some extent—arbitrary. To the best of our knowledge, no unified method combines—and assigns appropriate weight to—multiple lifestyle outcomes in

a way that optimally reflects its effect on cognitive aging. Therefore, we combined 6 lifestyle areas that are known to have an effect on cognitive aging, and we constructed the overall lifestyle score using lifestyle area–specific scores that reflect their respective effects on brain aging. A similar approach has been used previously in other fields [42,53,54]. Combining the self-reported lifestyle into one integer may result in some degree of misclassification and underestimation of the effect on lifestyle. However, in order to create the opportunity for meaningful understanding and also facilitate further analysis between lifestyle and long-term effects (eg, on cognition) we chose to create a single-digit lifestyle score, which can also be used as a starting point to better understand the magnitude of effects of multiple risk factors as well as to facilitate communication on the complete lifestyle effect as a whole.

Unanswered Questions and Future Research

Asking healthy people in the general population to participate in a lifestyle-improvement trajectory by following a predetermined intervention route that can last for a year (or longer) is notoriously difficult. Our most important recommendation for implementing an eHealth intervention and maximizing program adherence is to design public health programs that are highly flexible, enabling the participants to enter and exit the program freely and at their own convenience while still providing measurements at regular intervals. The ability to measure each participant's success using a more flexible approach should be addressed in future studies. Methods should be developed to identify pre–follow-up dropouts, who reached a sort of self-aspired end state, as those who successfully completed the intervention. Adapted stepped wedge cluster randomized trial designs may be well suited to this purpose [55]. From a more practical perspective, long-term adherence might be increased by "gamifying" future intervention programs. Adding game components to scientific research might provide the participant commitment and loyalty that many eHealth interventions currently lack [56]. The gamification of eHealth interventions can make the interventions more social, create competition, incorporate a reward system, and enhance motivation. Thus, within the constraints of playing a game, the working mechanisms of current eHealth components can be implemented [57].

Conclusions

In conclusion, we report adherence to and effectiveness of the BAM program, and we report that this Internet-based, self-motivated and self-managed eHealth intervention, which is aimed at changing multiple health behavior risk factors for cognitive decline, has a positive effect on public health. Given that the participants with the most room for improvement had the greatest change in behavior, and given that the participants who were the most involved with the program had the greatest benefit, we feel that future research with this tool is warranted. More globally, eHealth interventions can achieve more effective and more widespread primary prevention of cognitive decline, thus reducing the predicted strain of aging-related cognitive decline on health care systems in the near future.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Room for improvement in the Netherlands.

[PDF File (Adobe PDF File), 367KB - jmir_v18i6e171_app1.pdf]

Multimedia Appendix 2

Construction of the overall lifestyle score.

[PDF File (Adobe PDF File), 267KB - jmir_v18i6e171_app2.pdf]

References

1. Goldstein LB. Physical activity and the risk of stroke. Expert Rev Neurother 2010 Aug;10(8):1263-1265. [doi: 10.1586/ern.10.90] [Medline: 20662751]
2. Artaud F, Dugravot A, Sabia S, Singh-Manoux A, Tzourio C, Elbaz A. Unhealthy behaviours and disability in older adults: Three-City Dijon cohort study. BMJ 2013 Jul 23;347:f4240. [doi: 10.1136/bmj.f4240]
3. Elwood P, Galante J, Pickering J, Palmer S, Bayer A, Ben-Shlomo Y, et al. Healthy lifestyles reduce the incidence of chronic diseases and dementia: evidence from the Caerphilly cohort study. PLoS One 2013;8(12):e81877 [FREE Full text] [doi: 10.1371/journal.pone.0081877] [Medline: 24349147]
4. van Dam RM, Li T, Spiegelman D, Franco OH, Hu FB. Combined impact of lifestyle factors on mortality: prospective cohort study in US women. BMJ 2008;337:a1440 [FREE Full text] [Medline: 18796495]
5. Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Actual causes of death in the United States, 2000. JAMA 2004 Mar 10;291(10):1238-1245. [doi: 10.1001/jama.291.10.1238] [Medline: 15010446]
6. Mokdad A, Marks J, Stroup D, Gerberding J. Correction: actual causes of death in the United States, 2000. JAMA 2005 Jan 19;293(3):293-294. [doi: 10.1001/jama.293.3.293] [Medline: 15657315]

7.   Barnes DE, Yaffe K. The projected effect of risk factor reduction on Alzheimer's disease prevalence. Lancet Neurol 2011 Sep;10(9):819-828 [FREE Full text] [doi: 10.1016/S1474-4422(11)70072-2] [Medline: 21775213]

8.   Evans DA, Morris MC, Rajan KB. Vitamin E, memantine, and Alzheimer disease. JAMA 2014 Jan 1;311(1):29-30. [doi: 10.1001/jama.2013.282835] [Medline: 24381966]

9.   Solomon A, Mangialasche F, Richard E, Andrieu S, Bennett DA, Breteler M, et al. Advances in the prevention of Alzheimer's disease and dementia. J Intern Med 2014 Mar;275(3):229-250 [FREE Full text] [doi: 10.1111/joim.12178] [Medline: 24605807]

10. Anstey KJ, Mack HA, Cherbuin N. Alcohol consumption as a risk factor for dementia and cognitive decline: meta-analysis of prospective studies. Am J Geriatr Psychiatry 2009 Jul;17(7):542-555. [doi: 10.1097/JGP.0b013e3181a2fd07] [Medline: 19546653]

11. Blondell S, Hammersley-Mather R, Veerman J. Does physical activity prevent cognitive decline and dementia?: A systematic review and meta-analysis of longitudinal studies. BMC Public Health 2014;14:510 [FREE Full text] [doi: 10.1186/1471-2458-14-510] [Medline: 24885250]

12. Di Marco LY, Marzo A, Muñoz-Ruiz M, Ikram M, Kivipelto M, Ruefenacht D, et al. Modifiable lifestyle factors in dementia: a systematic review of longitudinal observational cohort studies. J Alzheimers Dis 2014;42(1):119-135. [doi: 10.3233/JAD-132225] [Medline: 24799342]

13. Goel N, Rao H, Durmer JS, Dinges DF. Neurocognitive consequences of sleep deprivation. Semin Neurol 2009 Sep;29(4):320-339 [FREE Full text] [doi: 10.1055/s-0029-1237117] [Medline: 19742409]

14. Horton L, Duffy T, Hollins MC, Martin C. Comprehensive assessment of alcohol-related brain damage (ARBD): gap or chasm in the evidence? J Psychiatr Ment Health Nurs 2015 Feb;22(1):3-14. [doi: 10.1111/jpm.12156] [Medline: 24846298]

15. Jonsdottir IH, Nordlund A, Ellbin S, Ljung T, Glise K, Währborg P, et al. Cognitive impairment in patients with stress-related exhaustion. Stress 2013 Mar;16(2):181-190. [doi: 10.3109/10253890.2012.708950] [Medline: 22746338]

16. Polidori MC, Schulz R. Nutritional contributions to dementia prevention: main issues on antioxidant micronutrients. Genes Nutr 2014 Mar;9(2):382 [FREE Full text] [doi: 10.1007/s12263-013-0382-2] [Medline: 24535714]

17. Wolf O. Stress and memory in humans: twelve years of progress? Brain Res 2009 Oct 13;1293:142-154. [doi: 10.1016/j.brainres.2009.04.013] [Medline: 19376098]

18. Kiliaan A, Arnoldussen IA, Gustafson D. Adipokines: a link between obesity and dementia? Lancet Neurol 2014 Sep;13(9):913-923 [FREE Full text] [doi: 10.1016/S1474-4422(14)70085-7] [Medline: 25142458]

19. Norton S, Matthews F, Barnes D, Yaffe K, Brayne C. Potential for primary prevention of Alzheimer's disease: an analysis of population-based data. Lancet Neurol 2014 Aug;13(8):788-794. [doi: 10.1016/S1474-4422(14)70136-X] [Medline: 25030513]

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

21. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]

22. Ritterband LM, Tate DF. The science of internet interventions. Introduction. Ann Behav Med 2009 Aug;38(1):1-3. [doi: 10.1007/s12160-009-9132-5] [Medline: 19816750]

23. Lustria ML, Noar SM, Cortese J, Van Stee SK, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. J Health Commun 2013;18(9):1039-1069. [doi: 10.1080/10810730.2013.768727] [Medline: 23750972]

24. 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 [FREE Full text] [doi: 10.1016/j.ypmed.2010.06.004] [Medline: 20558196]

25. Aalbers T, Baars MA, Olde Rikkert MG, Kessels RP. Puzzling with online games (BAM-COG): reliability, validity, and feasibility of an online self-monitor for cognitive performance in aging adults. J Med Internet Res 2013;15(12):e270 [FREE Full text] [doi: 10.2196/jmir.2860] [Medline: 24300212]

26. Eurostat. CBS. 2012. Nederland Europees kampioen internettoegang URL: https://www.cbs.nl/nl-nl/nieuws/2012/24/nederland-europees-kampioen-internettoegang [accessed 2016-06-14] [WebCite Cache ID 6iFzOax37]

27. Schwarzer R, Jerusalem M. Generalized Self-Efficacy scale. In: Weinman J, Wright S, Johnston M, editors. Measures in health psychology: A user's portfolio. Causal and control beliefs. Windsor, UK: NFER-NELSON; 1995:35-37.

28. Kelly RB, Zyzanski SJ, Alemagno SA. Prediction of motivation and behavior change following health promotion: role of health beliefs, social support, and self-efficacy. Soc Sci Med 1991;32(3):311-320. [Medline: 2024141]

29. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988 Jun;54(6):1063-1070. [Medline: 3397865]

30. Tangney JP, Baumeister RF, Boone AL. High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. J Pers 2004 Apr;72(2):271-324. [Medline: 15016066]

31. Aalbers T, Baars M, Qin L, de LA, Kessels R, Olde Rikkert MG. Using an eHealth Intervention to Stimulate Health Behavior for the Prevention of Cognitive Decline in Dutch Adults: A Study Protocol for the Brain Aging Monitor. JMIR Res Protoc 2015;4(4):e130 [FREE Full text] [doi: 10.2196/resprot.4468] [Medline: 26554496]

32.   Kiresuk T, Sherman R. Goal attainment scaling: A general method for evaluating comprehensive community mental health programs. Community Ment Health J 1968 Dec;4(6):443-453. [doi: 10.1007/BF01530764] [Medline: 24185570]

33.   Webb T, 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]

34.   Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]

35.   Bandura A. Social cognitive theory of self-regulation. Organizational Behavior and Human Decision Processes 1991 Dec;50(2):248-287. [doi: 10.1016/0749-5978(91)90022-L]

36.   Bandura A. Health promotion from the perspective of social cognitive theory. Psychology & Health 1998 Jul;13(4):623-649. [doi: 10.1080/08870449808407422]

37.   Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [Medline: 10170434]

38.   Fishbein M, Ajzen I. Predicting and changing behavior: the reasoned action approach. New York: Psychology Press; 2010.

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

40.   Milat AJ, King L, Bauman AE, Redman S. The concept of scalability: increasing the scale and potential adoption of health promotion interventions into policy and practice. Health Promot Int 2013 Sep;28(3):285-298 [FREE Full text] [doi: 10.1093/heapro/dar097] [Medline: 22241853]

41.   Elbert N, van Os-Medendorp H, van Renselaar W, Ekeland A, Hakkaart-van Roijen L, Raat H, et al. Effectiveness and cost-effectiveness of ehealth interventions in somatic diseases: a systematic review of systematic reviews and meta-analyses. J Med Internet Res 2014;16(4):e110 [FREE Full text] [doi: 10.2196/jmir.2790] [Medline: 24739471]

42.   Kirkegaard H, Johnsen N, Christensen J, Frederiksen K, Overvad K, Tjønneland A. Association of adherence to lifestyle recommendations and risk of colorectal cancer: a prospective Danish cohort study. BMJ 2010;341:c5504 [FREE Full text] [Medline: 20978063]

43.   Wieland L, Falzon L, Sciamanna C, Trudeau K, Brodney S, Schwartz J, et al. Interactive computer-based interventions for weight loss or weight maintenance in overweight or obese people. Cochrane Database Syst Rev 2012(8):CD007675 [FREE Full text] [doi: 10.1002/14651858.CD007675.pub2] [Medline: 22895964]

44.   Shephard RJ. Employee health and fitness: the state of the art. Prev Med 1983 Sep;12(5):644-653. [Medline: 6419221]

45.   Sweet SN, Fortier MS. Improving physical activity and dietary behaviours with single or multiple health behaviour interventions? A synthesis of meta-analyses and reviews. Int J Environ Res Public Health 2010 Apr;7(4):1720-1743 [FREE Full text] [doi: 10.3390/ijerph7041720] [Medline: 20617056]

46.   Nigg CR, Long CR. A systematic review of single health behavior change interventions vs. multiple health behavior change interventions among older adults. Transl Behav Med 2012 Jun;2(2):163-179 [FREE Full text] [doi: 10.1007/s13142-012-0130-y] [Medline: 24073109]

47.   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]

48.   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]

49.   Baker TB, Gustafson DH, Shah D. How can research keep up with eHealth? Ten strategies for increasing the timeliness and usefulness of eHealth research. J Med Internet Res 2014;16(2):e36 [FREE Full text] [doi: 10.2196/jmir.2925] [Medline: 24554442]

50.   Kreps GL, Neuhauser L. New directions in eHealth communication: opportunities and challenges. Patient Educ Couns 2010 Mar;78(3):329-336. [doi: 10.1016/j.pec.2010.01.013] [Medline: 20202779]

51.   Baker T, Gustafson D, Shaw B, Hawkins R, Pingree S, Roberts L, et al. Relevance of CONSORT reporting criteria for research on eHealth interventions. Patient Educ Couns 2010 Dec;81 Suppl:S77-S86 [FREE Full text] [doi: 10.1016/j.pec.2010.07.040] [Medline: 20843621]

52.   Pursey K, Burrows TL, Stanwell P, Collins CE. How accurate is web-based self-reported height, weight, and body mass index in young adults? J Med Internet Res 2014;16(1):e4 [FREE Full text] [doi: 10.2196/jmir.2909] [Medline: 24398335]

53.   Parekh S, Vandelanotte C, King D, Boyle F. Improving diet, physical activity and other lifestyle behaviours using computer-tailored advice in general practice: a randomised controlled trial. Int J Behav Nutr Phys Act 2012;9:108 [FREE Full text] [doi: 10.1186/1479-5868-9-108] [Medline: 22963354]

54.   Yun J, Won S, Kimm H, Jee S. Effects of a combined lifestyle score on 10-year mortality in Korean men and women: a prospective cohort study. BMC Public Health 2012;12:673 [FREE Full text] [doi: 10.1186/1471-2458-12-673] [Medline: 22900981]

55.   Mdege ND, Man M, Taylor NB, Torgerson DJ. Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation. J Clin Epidemiol 2011 Sep;64(9):936-948. [doi: 10.1016/j.jclinepi.2010.12.003] [Medline: 21411284]

56. Baranowski T, Buday R, Thompson D, Lyons EJ, Lu AS, Baranowski J. Developing Games for Health Behavior Change: Getting Started. Games Health J 2013 Aug;2(4):183-190 [FREE Full text] [doi: 10.1089/g4h.2013.0048] [Medline: 24443708]

57. Miller A, Cafazzo J, Seto E. A game plan: Gamification design principles in mHealth applications for chronic disease management. Health Informatics J 2016 Jun;22(2):184-193. [doi: 10.1177/1460458214537511] [Medline: 24986104]

Abbreviations

ANCOVA: analysis of covariance
BAM: Brain Aging Monitor
BAM-COG: Brain Aging Monitor–Cognitive Assessment Battery
BMI: body mass index
GAS: Goal Attainment Scaling
LTMS: long-term monitoring system
SD: standard deviation

Original Paper

If You Build It, Will They Come? Patterns of Internet-Based and Face-To-Face Participation in a Parenting Program for Military Families

Jennifer L Doty[1], PhD; Jessie H Rudi[2], PhD; Keri L M Pinna[3], PhD; Sheila K Hanson[4], PhD; Abigail H Gewirtz[2,5], PhD

[1]University of Minnesota, Division of General Pediatrics & Adolescent Health, Department of Pediatrics, Minneapolis, MN, United States

[2]Institute for Translational Research in Children's Mental Health, Minneapolis, MN, United States

[3]St. Catherine University, Psychology Department, St. Paul, MN, United States

[4]University of North Dakota, College of Business and Public Administration, Grand Forks, ND, United States

[5]University of Minnesota, Family Social Science, St. Paul, MN, United States

Corresponding Author:
Abigail H Gewirtz, PhD
University of Minnesota
Family Social Science
Room 294 McNH
1985 Buford Ave
St. Paul, MN, 55108
United States
Phone: 1 612 624 1475
Fax: 1 612 625 4227
Email: agewirtz@umn.edu

Abstract

Background: Some evidence suggests parents are drawn to media-based interventions over face-to-face interventions, but little is known about the factors associated with parents' use of Internet-based or Internet-enhanced programs, especially among military families. Research is needed to understand characteristics of parents who may be most likely to use online components or attend face-to-face meetings in order to ensure maximum engagement.

Objective: In this study, we examined characteristics that predict various patterns of Internet use and face-to-face attendance in a parenting program designed for military families.

Methods: An ecological framework guided analysis of differences in patterns of Internet-based use and face-to-face attendance by parents' demographic characteristics (gender, education, employment, and child age), incentives offered, and number of months the parent was deployed. We reported differences in the total number of online components completed over the 14 modules, total number of face-to-face sessions attended, and the use of different types of online components accessed (videos, downloadable handouts, mindfulness exercises, knowledge checks, and downloadable summaries). Then, we computed multinomial logistic regression accounting for nestedness (parents within families) to examine associations between demographic, programmatic, and military-related characteristics and patterns of engagement (use of online components and attendance at face-to-face sessions).

Results: Just over half (52.2%, 193/370) of the participants used the online components at least once, and the majority of participants (73.2%, 271/370) attended at least 1 face-to-face session. An examination of different patterns of participation revealed that compared with those who participated primarily in face-to-face sessions, parents who participated online but had little face-to-face participation were more likely to have received incentives than those who did not (95% CI 1.9-129.7). Among participants who had been deployed, those who had earned a 4-year degree (95% CI 1.0-2.2) and those who had been offered incentives to participate online (95% CI 2.1-58.6) were more likely to be highly engaged in online components and attend face-to-face compared with those who attended primarily face-to-face. However, those with a high number of months of deployment (95% CI 0.6-1.0) were less likely to be in the pattern of highly engaged in online components and face-to-face attendance. Compared with those who participated primarily face-to-face, deployed mothers were about 4 times more likely to engage in moderate online use with face-to-face attendance than deployed fathers (95% CI 1.21-11.83) and participate primarily online (95% CI 0.77-25.20).

XSL•FO
RenderX

Conclusions: Results imply that parents may be drawn to different delivery options of a parenting program (online components vs face-to-face sessions) depending on their education level, incentives to engage in online components, and their military-related experience. Results suggest potential directions for tailoring Internet-based interventions.

Introduction

Evidence-based parenting interventions have been shown to improve the well-being of both children and parents and have the potential to enhance family resilience in the face of chronic and acute stress [1-3]. However, the reach of these interventions is limited because of factors such as lack of trained practitioners and low parent participation [4-6]. These challenges have led researchers to begin examining parents' preferences for online program delivery formats [7,8], which hold the promise for reaching far broader swaths of the population. Online program components (eg, videos that demonstrate key skills and downloadable handouts) are potentially an important supplement to traditional, face-to-face interventions [6,9]. However, we know little about which parents most engage with these program components.

More specifically, in military families, participation in interventions may be affected by the deployment of parents. A recent study revealed that more than half of active-duty military families who initially engaged in a prevention program did not complete it because of work-related issues or deployments [1]. Reserve component (ie, National Guard and Reserves; NGR) service members participate in monthly weekend drills and extended annual trainings in addition to their civilian work commitments. National Guard and Reserve populations also have been extensively deployed in the recent conflicts in Iraq and Afghanistan. There is a dearth of literature investigating the potential barriers to participation in parenting skills programs designed for military families, and particularly NGR families.

In this population, knowledge is also lacking regarding the relationship between demographic characteristics, ecological characteristics (eg, number of deployments), and parents' use of online components in intervention settings. This information will allow service providers to tailor evidence-based programs for online distribution and maximize access to beneficial programs for military families. Online program components may be particularly beneficial to offset potential barriers to in-person participation. Understanding demographics and other characteristics of parents who engage in online components to differing degrees can inform resource allocation, recruitment, and retention efforts. To extend previous research on parent engagement in Internet-enhanced parenting programs, this study examined how demographic, programmatic, and military deployment-related factors are associated with patterns of parents' use of online parenting components and face-to-face attendance in a parenting program for military parents.

Engaging Parents Online

Online Parenting Programs

A growing number of Internet-based interventions have been geared specifically toward parents and parenting [10,11]. One study examined the efficacy of an 8-module, intensive, positive parenting program, Triple P Online, for parents of children with early-onset disruptive behavior problems using a randomized controlled trial (RCT) design [12]. Parents receiving the Internet-based intervention had significantly better outcomes on measures of children's problem behavior, dysfunctional parenting styles, parents' confidence in their parenting role, and parental anger compared with the control group (general Internet use). These gains were generally maintained, and in some cases enhanced, at the 6-month follow-up assessment. Additionally, satisfaction ratings for the Internet-based intervention program were high [12].

Similarly, another RCT evaluated the efficacy of an Internet-based parent management training program for children with conduct problems [13]. Children whose parents had participated in the Internet-based program showed a greater reduction in conduct problems than children in the waitlist control group. Parents who completed the Internet-based parenting program also reported less use of harsh and inconsistent discipline after completing the program, as well as more positive praise. These positive effects were maintained at the 6-month follow-up. Other RCTs have also found that Internet-based parent management training programs are effective in reaching parents and report encouraging results [10,14-17].

Although these randomized trials are important for establishing the efficacy of Internet-based programs, they revealed very little about which parents might gravitate toward a face-to-face program and which might gravitate toward online materials because all intervention families were assigned to the Internet-based treatment. Often, Internet-based programs are viewed as a way to reduce barriers to engagement, but research comparing in-person versus Internet-based programs (ie, comparative effectiveness trials) is lacking. Studies that include Internet-based programming as a supplement to a face-to-face program can begin to answer some of these questions about differences in parents' online usage. Sanders and colleagues [18] found that online supplemental materials were as effective as self-help workbooks for parents. In addition, Internet-based supplements to a parenting prevention program presented via television were found to be an effective method of engaging hard-to-reach individuals [19]. Little is known, however, about what contextual factors—individual, familial, or ecological—are

associated with parents' use of Internet-based programs and supplements compared with face-to-face programming.

Parents' Preferences

Efforts to understand how parents engage in parenting programs have used marketing methods to explore preferences for program delivery options [7]. Cunningham and colleagues [7] identified subgroups of parents who preferred different resource delivery formats. The "Action" segment of parents (ie, those highly motivated to engage in services) preferred weekly, in-person meetings with other parents in addition to coaching phone calls from a therapist, whereas the "Information" segment of parents preferred to solely receive parenting information without in-person meetings or support [7]. The "Overwhelmed" segment of parents was more likely to have a child with externalizing problems but was less likely to prefer information or professional support than other segments of parents. Interestingly, however, overwhelmed parents reported preferring information found on the Internet significantly more than those in the other two segments. Metzler and colleagues [8] found a preference among parents for media-based programs (eg, television, online) rather than face-to-face programs in a survey of 162 parents with children aged 3-6 years. Overall, parents reported preferring television programs, online formats, and written materials to face-to-face interventions. Program materials that are tailored to parents' individual preferences may improve engagement in and adherence to Internet-based parenting programs [7,20].

The studies discussed above examined parent preferences, rather than actual behavior. Moreover, some evidence suggests that parents' preferences do not always match actual engagement in various methods of program delivery, and more information about parents' behavior in practice is needed [7]. This study fills a gap in the literature by examining the use of Internet-based programming and patterns of face-to-face attendance in a parenting program for military families and the characteristics and contexts that predict those patterns.

Demographic Characteristics Associated With Online Use

Although little is known regarding parent characteristics related to adherence to Internet-based prevention programs, parent and family demographic variables have been shown to be related to parents' general use of the Internet and other technology. Specifically, income and education were associated with parents' online information-seeking behaviors [21-23]. In another study, parents who were on a waitlist for mental health services for their children and preferred Internet support had higher levels of educational attainment compared with those who preferred group or face-to-face interim services [24]. Although seeking information online differs from participating in an Internet-based parenting program, these studies help inform our understanding of parents' use of Internet-based programs.

Gender is another important demographic characteristic to consider. Although some research has found that women do not use the Internet as frequently as men [25], other studies have found no gender differences in Internet use [26]. However, mothers have tended to use more parenting content online than

fathers [27,28]. A study of parents' preferences for online, parenting video episodes found that mothers tended to rate the online modules as more engaging, watchable, and realistic than fathers did [8]. Mothers may find parenting content delivered via media, such as television and the Internet, more interesting and relevant compared with fathers. Existing parenting content may be tailored more to mothers' unique needs than fathers' [29]. Alternatively, mothers may be gatekeepers to fathers' access to parenting content [30] or fathers may lack parenting efficacy given the traditional role mothers have played in child-rearing [31]. Together, the aforementioned studies suggest that parents' demographic characteristics are linked to online preferences and use.

Face-To-Face Attendance

The greater body of research on engaging parents in face-to-face parenting programs may inform research on Internet-based program participation. Face-to-face attendance in parenting programs varies widely; studies report that parents attend 35% to 61% of face-to-face group sessions [4,32,33]. A meta-analysis found a small but significant negative effect of socioeconomic status (SES) on dropout rates [34], addressing prior research that had reported mixed findings in this regard [35-37]. Similarly, single parents have been found to attend fewer face-to-face sessions than married or cohabiting parents in some studies [38] but not in others [32,39]. Family size is also related to parents' face-to-face attendance in parenting programs, such that those with large families tend to attend fewer face-to-face sessions [32]. Overall, the evidence suggests that families at slightly higher risk (ie, single-parent families, lower SES families, and those with more children), who arguably may benefit most from such programs, may actually attend fewer sessions and participate less [32,38,40].

Evidence also suggests that process variables within the prevention program experience (eg, interest in and comfort with the group) may be related to continued participation [40]. For example, Fox and Gottfredson [41] found that parents who completed the program tended to report higher interest and more comfort with the pretest or program participation than parents who did not complete the program. In another study, interaction with other participants and group dynamics contributed to retention [40].

Although parents report that small incentives would promote attendance [42], studies have had mixed results regarding the effectiveness of providing incentives to encourage attendance. Al-Halabí Díaz and Errasti Pérez[43] found that the use of small incentives at the end of each session improved face-to-face attendance. Similarly, other studies found that incentives were positively associated with face-to-face attendance [44,45]. However, a relationship between incentives and attendance has not been found in other studies [46,47]. Although much has been written about attendance in face-to-face parenting programs, much less is known about participation in online settings.

Ecological Characteristics

From an ecological systems framework, families' ability to participate in parenting programs may be influenced by barriers

or facilitators in their immediate environment [48,49]. On the basis of the ecological assumption that basic resources in the environment sustain families and communities [48], those who most need support may also be those who lack the resources to access that support. However, while the aforementioned literature outlines barriers for other at-risk populations, very little is known about the barriers to participating in parenting programs for military families.

Among active duty military families, levels of practical and emotional support may be high because of a sense of community found on military bases [50]. However, NGR families experience more isolation, as they are geographically dispersed across the United States without access to resources found on military installations. National Guard and Reserve parents may also feel especially pressed for time as many have civil employment during the week and military training and duties on weekends [1], and this may be a substantial barrier to face-to-face participation in parenting programs.

Internet-based resources may provide support to parents who feel isolated as a result of a partner's deployment. Although technology has become increasingly accessible [51], little is known about which members of military families are most likely to engage with Internet-based resources. Evidence suggests that the deployment cycle in particular may increase families' use of technology to stay in touch [52,53]. During deployment, parents who are on the home front, the majority of whom are mothers, are accustomed to the use of Internet-based technology for communication with deployed service members. Research also suggests Internet-based programming may be well suited for military families post deployment [54]. However, those who have extensive prior deployments may be less likely to engage in parenting programming if they are struggling with multiple demands to reenter civilian life or posttraumatic stress symptoms [55].

Our Study—After Deployment, Adaptive Parenting Tools

In our study, we sought to identify patterns of Internet-based and face-to-face participation in After Deployment, Adaptive Parenting Tools (ADAPT), an Internet-enhanced parenting program for military families. We examined the demographic, programmatic, and military-related characteristics associated with different patterns of engagement. Understanding patterns of engagement is important because the success of prevention programs and the benefits parents receive from these programs depend on their ability to engage and retain parents. This study analyzed data from an RCT of ADAPT, which aimed to improve parenting and child adjustment in families after the reintegration of a parent from deployment to Iraq or Afghanistan. The ADAPT program is a 14-week parent training program that provided online components as supplements to weekly face-to-face programming.

To understand patterns of engagement with online components and face-to-face participation with military parents, we sought to answer the following research question: How often did parents use various online components and attend face-to-face sessions? Furthermore, based on the ecological systems framework, we hypothesized the following:

- H_1: Characteristics of parents (higher income, more education, full or part-time employment, fewer children, younger children, married, and female) and the program (incentives) will be positively associated with patterns of greater online and face-to-face use.
- H_2: Characteristics that may pose a barrier to military families' participation (number of deployments) will be positively associated with patterns of lesser online and face-to-face use.

Methods

Procedures

This study includes the subset of families from the larger study (N=336 families) who were randomly assigned to the treatment group—that is, to participate in a preventive intervention program, ADAPT (n=207 families; n=370 parents). Families were eligible to participate in the RCT if at least one parent had been deployed overseas in service of Operation Enduring Freedom (Afghanistan), Operation Iraqi Freedom (Iraq), or Operation New Dawn (Iraq) since 2001, if the family had at least one child between the ages of 4 and 12 years, and if they were willing to participate in the parenting program if invited. Participants were recruited in several ways: project staff presence and presentations at military-sponsored events (eg, reintegration events and military family picnics), referral from military personnel (eg, family readiness group leader or commander), word of mouth from fellow service members or another parent, and media (eg, television or radio, advertisements, and online social media). Families could go directly to the ADAPT website to consent to participate or request to be contacted for more information. Recruitment staff replied via phone call to answer questions, emailing the hyperlink for the screener and online consent form as needed.

In two-parent families, both parents were invited to participate in the online and in-home baseline assessments after individually completing the online consent form. After baseline assessments, families were randomized to either the ADAPT parenting program or the control group (services-as-usual). Staff then contacted the family to inform the parents of the result of randomization and discuss arrangements for those who were randomized to ADAPT to attend group sessions. Specifically, groups were arranged to be delivered in parents' geographic area, as close to their home as possible. If no groups were currently occurring in their geographic area, parents were invited to the next available group after their baseline assessment. As these arrangements were made for determining which geographic group a family would attend, staff also encouraged both parents in two-parent families to attend the program. Staff explained to such families that although attendance by both parents was expected to be most beneficial, it was not required (eg, if one parent had to work, the other parent could attend alone). A total of 7 cohorts of groups were delivered, each cohort included 2 to 7 groups, and each group included 3 to 10 families (up to 16 individual parents per group).

After Deployment, Adaptive Parenting Tools Program

The ADAPT program is an adaptation of Parenting Through Change, a 14-week group-based Parent Management Training–Oregon Model (PMTO) preventive intervention, for military families [56]. In addition to targeting positive parenting practices that are core to PMTO interventions (skill encouragement, positive involvement, family problem solving, monitoring, and effective discipline), ADAPT provides intensive coaching for parents on emotion regulation (via mindfulness exercises) and emotion coaching skills [57]. Skills are taught in 14 two-hour weekly group sessions using active teaching methods such as role-playing and group discussions. For more information about the ADAPT program, see Gewirtz and colleagues [54,57].

Online ADAPT Components

Building upon previous research demonstrating the many barriers to face-to-face attendance at prevention programs [5] as well as difficulties with retention [37,44], an Internet-based supplement was developed to engage parents with the program content in their own homes as much as possible (see Multimedia Appendix 1 for example screenshots). The 56 online components were supplemental materials organized by the 14 group sessions into 14 online modules. Each module included a menu of components such as access to skill and practice videos of military families who are learning and practicing key parenting skills with their children, audio recordings of mindfulness exercises, knowledge checks, and printable PDF documents summarizing key parenting skills. Parents were able to choose the components that best fit their needs for a personalized approach to the online materials. The number of components in the modules ranged from one in module 14 to eight in module 5. After each group session, parents received an email prompt directing them to the relevant online module for that week. ADAPT facilitators encouraged parents to view online material between sessions and to share the material with family members. Parents who were unable to attend face-to-face sessions were given the option to complete the online program modules when delivery of group sessions was complete (ie, after all possible face-to-face attendance options had been exhausted).

Online and Face-To-Face Program Incentives

In an effort to encourage online engagement after initial low participation online, those in cohorts 3 through 7 who used the online components the previous week were entered into a drawing for a US $25 gift card at the face-to-face group session. Families participating in face-to-face sessions received a US $15 gift card for attendance at each session to compensate them for their travel and time. Participants who did not attend face-to-face sessions but were offered to complete the program online in the last cohort were provided a US $15 gift card at the completion of each online session.

Measures

Use of Online Components

Parents' use of the online components was recorded using an online data tracking system. The system recorded whether each parent accessed an online component via clicking (coded as 0=did not click, 1=click). Clicks were interpreted as parents accessing and using that online component, and date stamps showed that parents accessed modules consecutively on different days, an indication that they were not casually browsing the material at one sitting. The total number of all components used was calculated as well as the total number of each component type used (videos, mindfulness exercises, knowledge checks, and downloadable summaries). In an effort to maximize use of online components, incentives were provided to families in cohorts 3 through 7 and in the online cohort (0= *did not receive incentives to go online*; 1= *received incentives to go online*).

Face-To-Face Session Attendance

Participants' attendance at face-to-face sessions was recorded using sign-in sheets, receipts from gift card payments, and facilitator records. Number of sessions attended was summed to create a variable indicating the number of face-to-face sessions attended (of those who attended at least one session: mean 8.50, SD 3.91; including those who did not attend: mean 6.26, SD 5.04).

Variables Relevant to Military Families

Parents who had served in the military reported the number of months they had been deployed in the recent conflicts since 2001. A dichotomous variable reflecting deployment was created for the entire sample (0= *not deployed*; 1= *deployed in recent conflicts*).

Demographics

Parents self-reported demographic characteristics at their initial in-home assessment. Socioeconomic variables included household income in US dollars (less than $10,000 per year; $10,000 or more per year in $10,000 increments up to $150,000; or more than $150,000 per year) and education level (some high school or less, General Educational Development test, high school diploma, some college, associate's degree, 4-year college degree, master's degree, or doctoral or professional degree). Parents reported the number of children living at home at least 50 percent of the time, target child's date of birth (converted to age), marital status (dichotomized into married and not married), and employment status (dichotomized into employed full-time or part-time and not employed).

Data Analysis

We addressed the research question regarding the frequency of parents' online component use and face-to-face attendance using descriptive analyses (t tests and correlations). We also examined the use of different types of online components accessed, including videos, mindfulness exercises, knowledge checks, and summaries using descriptive analyses. We then identified patterns of use of online components and face-to-face attendance, and parents were categorized into descriptive groups based on their frequency of participation online and face-to-face. To test the hypotheses that demographic, programmatic, and military-related variables would predict patterns of use, multinomial logistic regression with standard errors adjusted to account for the nested data structure (ie, parents within families) was computed in Stata 14 [58]. We examined associations between demographic characteristics (education, income, employment, number of children, marital status, child age, and gender), cohort (incentive to participate online), and

military-related characteristics (months deployed) and patterns of online and face-to-face participation. Because deployment and gender were confounded (almost all men were deployed), additional analyses included only those who had been deployed.

Missing data analysis revealed very little missing data. Only 2 variables had missing data that made up more than one percent of the total sample: number of children (1.9%) and income (1.4%). Therefore, listwise deletion was used, resulting in an analytic sample of 365 individuals in the main multinomial logistic regression model.

Results

Descriptive Statistics

Of the 207 invited families, 84.5% (175/207) participated in at least one of either the online modules or the face-to-face sessions. A total of 75.4% (156/207) participated face-to-face and a total of 68.6% (142/207) participated online (9.2%, 19/207, participated only online). Of the participating individuals, 88.3% (325/368) were married, and on average, most parents reported having 2 or more children. Parents were asked to choose their youngest child between the ages of 4 and 12 years to be the primary focus child for the assessments. The average age of families' target child was 7.7 years. Parents reported annual household income as follows: 25.5% (93/365) had an annual income of less than US $50,000; 43.3% (158/365) had an annual income between US $50,000 and US $99,999; and 31.2% (114/365) had an income of more than US $100,000.

We also analyzed parents' engagement in the program as individuals: 78.9% (292/370) participated in either the online modules or the face-to-face sessions. A total of 73.2% (271/370) participated face-to-face and a total of 52.2% (193/370) participated online (5.7% participated only online; 21/370). Approximately half of the participants were mothers (51.1%, 188/368), and 90.5% reported being European American.

Parents reported belonging to the following branches of the military: Army National Guard or Reserve (44.3%, 163/368), Air Force National Guard or Reserve (7.9%, 29/368), Navy Reserve (1.6%, 6/368), and other (6.8%, 25/368); 39.4% (145/368) were civilian partners. On average, the number of deployments for service members was 1.7, and the average number of months of deployment was between 13 and 18 months. Of the participants, 50.3% (185/368) had completed a 4-year degree or more and 67.1% (247/368) were employed full-time.

Online and Face-To-Face Participation

Descriptive analyses revealed that the overall pattern of use of online components was bimodal (ie, no use or high use; Figure 1). The percentages of mothers and fathers who accessed each component at least once were calculated (Table 1). Because the distribution was bimodal (as has been found in past trials of PMTO, those who participated at least 3 times were more likely to complete the program), a cutoff of 4 components was used to calculate both online and face-to-face use. Among participants who accessed at least 4 online components, the mean number of components accessed was 41. Online videos and handouts were accessed most frequently. The first components in the first modules had the heaviest use. Significant differences in online component use were found by gender, education, incentives, and past deployment (Table 2). Mothers, parents with a 4-year degree, parents who received incentives to participate online, and those who had not been deployed used more components compared with fathers, parents without a 4-year degree, parents who did not receive incentives to go online, and those who had been deployed. In Figure 2, a graph illustrates the rate of online participation in 3 phases of incentives to go online: phase 1, no incentives; phase 2, a drawing for those who went online; and phase 3, gift cards for those who went online. Correlations revealed a significant but trivial, positive relationship between online use and number of months deployed (.10, P=.050).

Table 1. Parents' use of online components (n=370).

Component	Total parents		Mothers		Fathers	
	n	%	n	%	n	%
Any online use	193	52.2	115	60.5	78	43.3
Videos	185	50.0	113	59.5	72	40.0
Knowledge checks	170	45.9	102	53.7	68	37.8
Mindfulness	176	47.6	109	57.4	67	37.2
Handouts	183	49.6	110	57.9	73	40.6

Face-to-face attendance was also bimodal (Figure 3). Among participants who attended at least 4 sessions, the average number of sessions attended was 9.8. Mothers attended marginally more than fathers (Table 2). Correlations revealed a statistically significant, negative relationship between face-to-face attendance and child age—that is, parents with younger children attended more sessions (−.14, P=.009).

Table 2. Results of *t* tests of participation in online components and attendance.

Online Components	Gender			Education			Incentives			Deployed		
	M[a]	F[a]	*P*	<4 years	4 years+	*P*	Y[a]	N[a]	*P*	Y	N	*P*
Total	17.0	23.2	.013	18.1	22.2	.104	23.0	8.4	<.001	18.2	22.7	.077
Videos	4.9	6.5	.007	5.2	6.2	.109	5.1	2.2	<.001	5.1	6.6	.023
Knowledge checks	1.8	2.6	.008	2.0	2.4	.088	2.5	0.9	<.001	2.0	2.5	.060
Handouts	3.7	4.7	.055	3.8	4.7	.095	4.9	1.8	<.001	3.9	4.7	.166
Mindfulness	3.9	5.4	.011	4.2	5.1	.155	5.4	1.4	<.001	4.2	5.3	.063
Attendance	5.8	6.7	.088	6.4	6.1	.560	—	—	—	6.08	6.49	.444

[a] M: male; F: female; Y: yes; N: no.

Guided by the distributions of online component use and face-to-face attendance (Figures 1 and 3), five patterns of program engagement were identified (Figure 4): (1) face-to-face attendance with high online use, (2) face-to-face attendance with moderate online use, (3) primarily face-to-face attendance, (4) primarily online use with little face-to-face attendance, and (5) little to no attendance or online use. The difference between high and moderate online use was determined by a median split of the number of online components used.

To test our hypotheses that demographic, programmatic, and ecological characteristics (ie, military related) would be associated with various patterns of participation, multinomial logistic regressions were computed, adjusting the standard error to account for individuals nested in families. The reference group of parents who attended primarily face-to-face was chosen because it represented traditional delivery. Because the vast majority of the men in this sample had been deployed, gender and deployment were confounded and modeled separately in preliminary models. The first hypothesis that demographic and programmatic characteristics would be associated with patterns of greater online and face-to-face use was partially supported. The following findings note significant differences with a *P* value of less than .05 compared with the reference group, parents who attended primarily face-to-face. The first model fit the data well (χ^2_{20}=72.8, *P*<.000), and gender was a significant predictor of each pattern (results available upon request). The second model, in which we removed gender but added deployment, also fit the data well (χ^2_{20}=62.9, *P*<.000), and deployment was a significant predictor of each pattern (see Tables 3 and 4). Those who had face-to-face attendance and high engagement online had higher levels of education. Those who were employed and who had older children were more likely to have a pattern of little face-to-face or online participation. Those who were incentivized to go online were about 2 to 9 times more likely to have a pattern of face-to-face attendance and high or moderate online use and about 15 times more likely to have a pattern of primarily online use. Income, number of children, and marital status were not associated with patterns and were therefore removed from the models for parsimony.

Table 3. Multinomial logistic regression of characteristics associated with patterns of online component use and face-to-face attendance (n=365).

Characteristic	Attendance + high online use (n=71)			Attendance + moderate online use (n=79)		
	RR[a]	*P*	95% CI	RR	*P*	95% CI
Deployed	0.44	.017	0.23-0.87	0.47	.027	0.24-0.92
Education	1.38	.025	1.04-1.84	1.21	.162	0.93-1.59
Employed	0.75	.518	0.32-1.79	2.17	.099	0.86-5.45
Child age	1.13	.094	0.98-1.29	1.10	.150	0.97-1.25
Incentives	9.69	.000	3.23-29.03	1.96	.086	0.91-4.23

[a]RR denotes relative risk. Reference group attended at least 4 times but had little to no online use (n=77).

Table 4. Multinomial logistic regression of characteristics associated with patterns of online component use and little face-to-face attendance (n=365).

Characteristic	Little/no attendance + online use (n=28)			Little/no attendance or online use (n=110)		
	RR[a]	P	95% CI	RR	P	95% CI
Deployed	0.45	.069	0.19-1.07	0.57	.049	0.32-1.00
Education	1.40	.107	0.93-2.12	1.11	.424	0.85-1.46
Employed	0.55	.274	0.19-1.60	3.10	.012	1.29-7.47
Child age	1.18	.064	0.99-1.41	1.21	.005	1.06-1.38
Incentives	15.94	.010	1.96-129.69	1.70	.184	0.78-3.70

[a]RR denotes relative risk. Reference group attended at least 4 times but had little to no online use (n=77).

The second hypothesis that military-related variables would be associated with patterns of lesser online and face-to-face use was partially supported. To avoid the confound between gender and deployment, only those men (n=173) and women (n=34) who had been deployed were selected in the final model and gender was added into the model (n=207). The final model fit the data well ($\chi^2=40.67_{24}$, P=.004). Compared with parents who attended primarily face-to-face, mothers who had been deployed were nearly 4 times more likely to have moderate online use and attend face-to-face and were more than 4 times more likely to attend primarily online, although, likely because of the relatively small number of women, the effect is only marginally significant (see Tables 5 and 6). Compared with parents who attended primarily online, those with fewer months deployment were less likely to have a pattern of high online use and face-to-face attendance.

Table 5. Multinomial logistic regression of characteristics associated with patterns of online component use and face-to-face attendance among those who had been deployed (n=207).

Characteristic	Attendance + high online use (n=34)			Attendance + moderate online use (n=43)		
	RR[a]	P	95% CI	RR	P	95% CI
Months deployed	0.78	.041	0.61-0.99	0.88	.265	0.71-1.10
Education	1.53	.029	1.04-2.24	1.31	.140	0.91-1.89
Employed	0.47	.334	0.10-2.18	0.77	.718	0.18-3.24
Child age	1.05	.617	0.87-1.26	1.01	.919	0.85-1.20
Female	1.15	.849	0.28-4.71	3.78	.023	1.21-11.83
Incentives	11.20	.004	2.14-58.60	2.80	.068	0.92-8.49

[a] RR denotes relative risk. Reference group attended at least 4 times but had little to no online use (n=55).

Table 6. Multinomial logistic regression of characteristics associated with patterns of online component use and little face-to-face attendance among those who had been deployed (n=207).

Characteristic	Little/no attendance + online use (n=13)			Little/no attendance or online use (n=66)		
	RR[a]	P	95% CI	RR	P	95% CI
Months deployed	0.92	.563	0.69-1.23	0.86	.164	0.69-1.06
Education	1.25	.402	0.74-2.09	1.16	.389	0.83-1.62
Employed	0.50	.503	0.64-3.86	0.69	.567	0.19-2.46
Child age	1.12	.394	0.87-1.43	1.16	.044	1.00-1.35
Gender	4.41	.095	0.77-25.20	1.00	.994	0.29-3.50
Incentives	8.69	.044	1.06-71.22	2.03	.113	0.84-4.92

[a]RR denotes relative risk. Reference group attended at least 4 times but had little to no online use (n=55).

Figure 1. Histogram of Online Components Participants Accessed (n=370).

Online Components
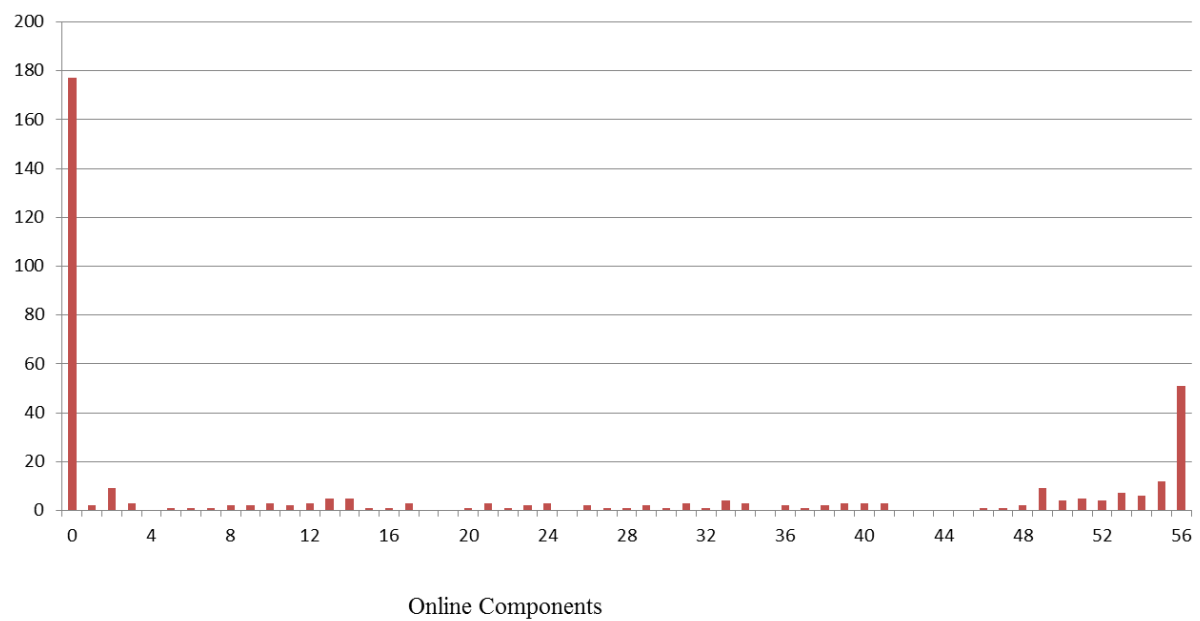
Figure 2. Graph of 3 phases of incentives for online participation (n=370).

XSL•FO
RenderX
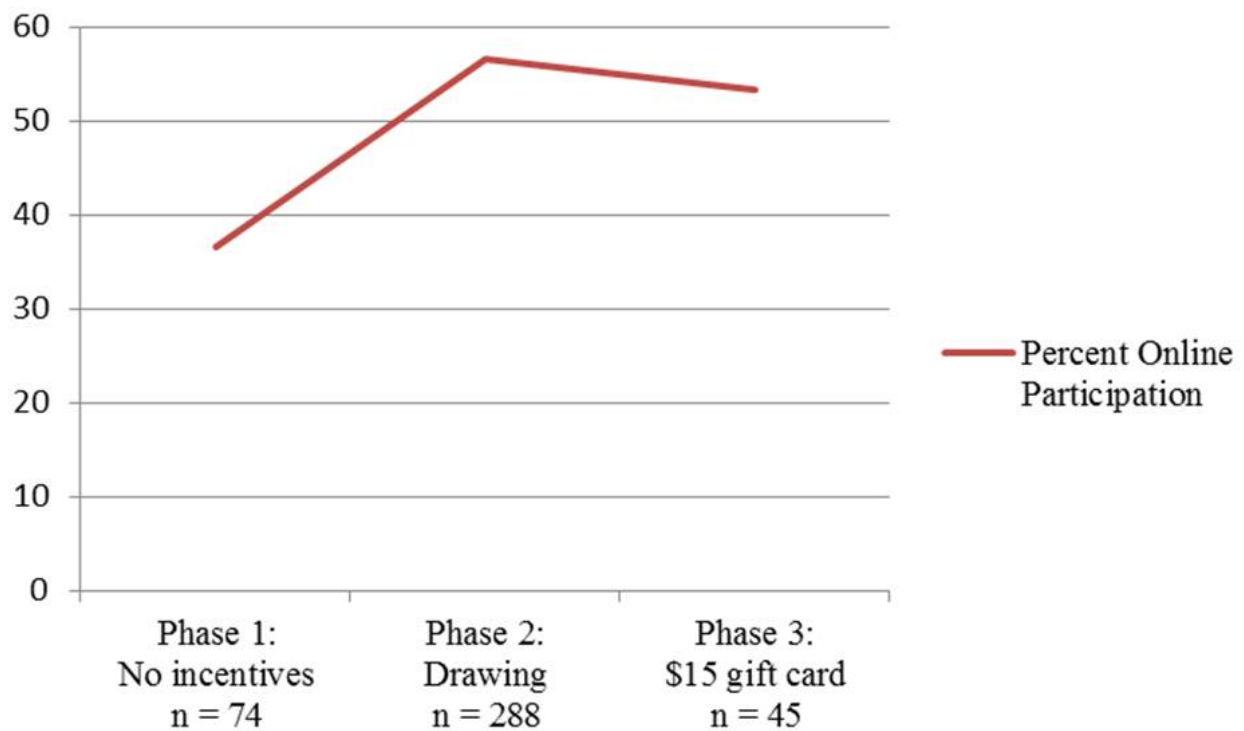
Figure 3. Histogram of Total Number of Face-to-Face Sessions Participants Attended (n=370).
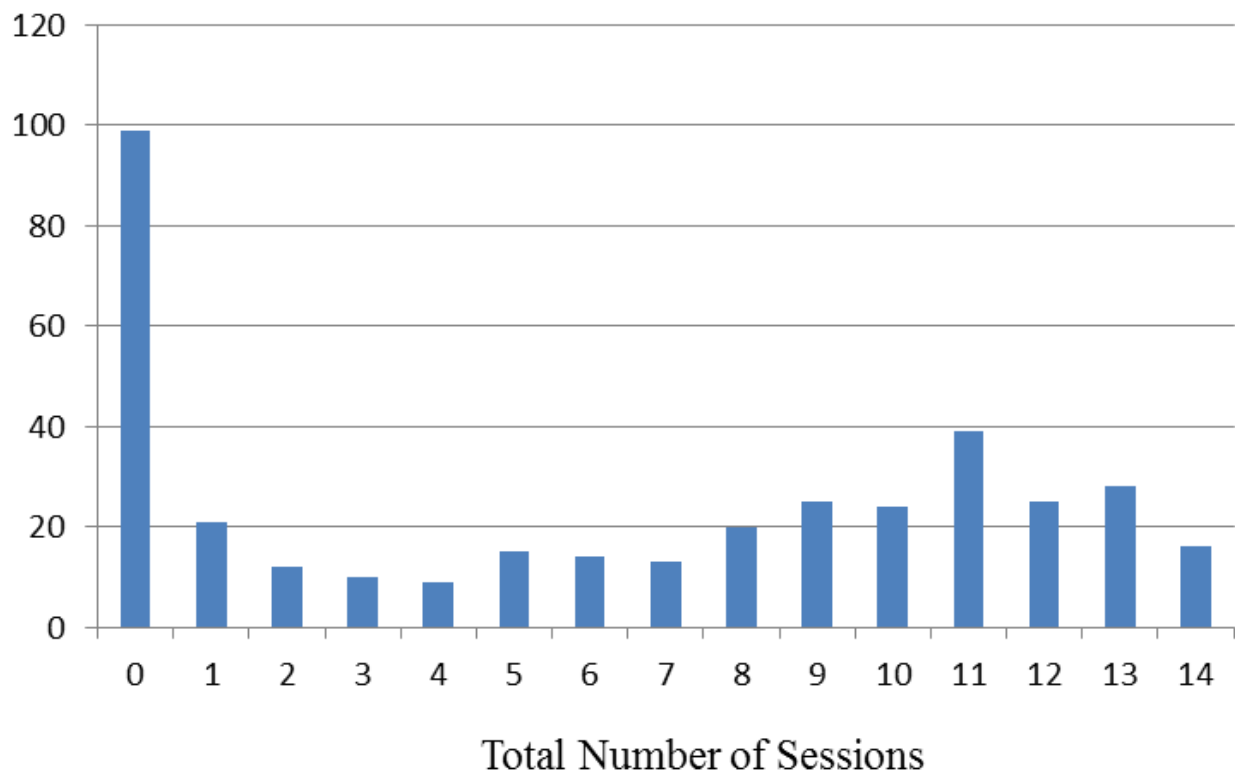
Figure 4. Patterns of online use and face-to-face attendance (analytic sample; n=365).
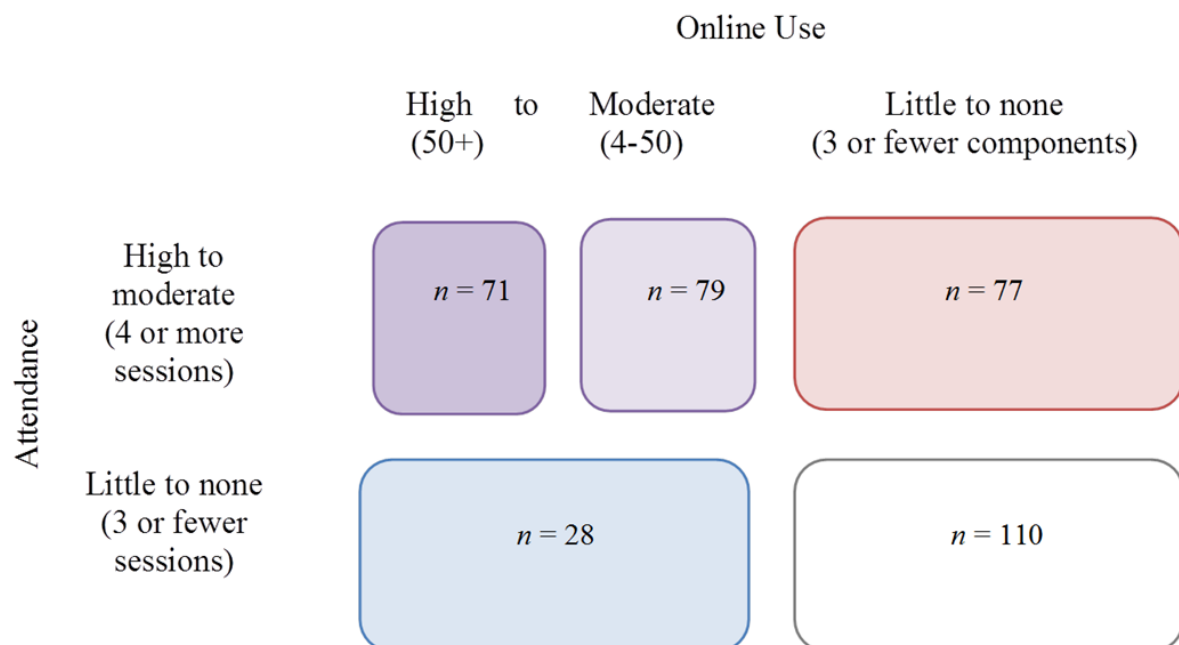
Discussion

To better understand patterns of engagement in a preventive parenting intervention program with both online components and face-to-face sessions, we sought to identify demographic, programmatic, and ecological characteristics of military life related to parents' use of online components and face-to-face attendance. As in prior studies evaluating PMTO (personal oral communication, Marion Forgatch, PhD, January, 2014), most participants had a high frequency of attendance (on average 10/14 sessions) once they participated in more than 3 face-to-face sessions. We also found this to apply to online components, such that parents who completed more than 3 online components had a high frequency of online component use (on average, 41/52 components). The high percentage of total participation (84.5% of families, 78.9% of individuals) was comparable to or better than other parenting programs [32,37]. Five patterns of engagement were identified: high online use with face-to-face attendance, moderate online use with face-to-face attendance, primarily face-to-face attendance, primarily online use, and little to no online use or attendance.

This study found that among those who had been deployed, participants with a high number of months of deployment were less likely to be in the group with the highest engagement both online and face-to-face. Although military families communicate via Internet-based medium during deployment [52,53], it may be that at home those who spent the most time away value face-to-face time with family rather than spending time on the computer. Alternatively, these individuals may be less connected to parenting, given their lengthy deployments and absences from the family, and thus see less value in participating in a parenting program. Employment was not significantly associated with participation among those who had been deployed, but it was significantly associated with lower participation across the entire sample, which suggests that employment of the nondeployed spouse may have been a barrier to participation.

Incentives seemed to be important motivators for parents to participate in online components; these findings parallel past findings regarding the use of incentives with face-to-face attendance [43,45]. Parents participating before the introduction of incentives (ie, phase 1) had low participation online, but an increase in online participation was evident in phases of delivery where incentives were offered to participate online. Although the proportion of those who participated online among those who were incentivized with a drawing in phase 2 was similar to those who were incentivized directly with a gift card in phase 3, it should be noted that those who were recruited in phase 3 were "hard to reach" participants who had not previously been able to attend face-to-face. Almost 10% of families who participated in ADAPT did so primarily online and likely would not have participated in the program at all without the online option. Overall, this evidence suggests that online options (especially those that are incentivized) could increase engagement in parenting programs and provide access to resources for parents who are isolated or unable to attend face-to-face programs.

Similar to previous research findings [27], mothers were more likely to engage with online parenting content than fathers. This study is the first to demonstrate that mothers who had been deployed were more likely to participate in an online setting compared with participation in the traditional face-to-face setting. Deployed mothers may be a particularly appropriate target population for online programs, given these findings, and their likely high use of online technologies during deployment.

We also found that those with the highest levels of education were the most engaged online. This finding corroborates past research that found that parents with higher income levels were more likely to prefer online services and seek parenting information online [21,24] than low-income parents. We also found that parents with a 4-year degree were significantly more likely to access knowledge checks (quizzes) and handouts but not video components or mindfulness exercises, compared with those without a 4-year degree. Although parents in lower socioeconomic circumstances increasingly have access to online content, past research has suggested that digital literacy deficits may continue for some individuals because they lack skills or time to process online content [22]. For example, less educated parents may not feel that they have the confidence to complete quizzes. These findings suggest that demographic characteristics are important to consider in engaging military families in online and face-to-face parenting programs.

Limitations and Future Directions

Although this study is novel in its examination of demographic, programmatic, and ecological characteristics of military life as they relate to participation in an Internet-enhanced parenting program, there are some limitations to consider. This study included a sample of Midwestern NGR families with a parent who had been deployed, the majority of whom were European American. Although our findings help us better understand program engagement in this important subset of parents, our results may not generalize to active duty military families or civilian families. In addition, past research has found barriers to participation among those with low SES, but this sample did not have high variability of SES among participants (ie, most were upper middle class), and therefore the results may not adequately reflect the socioeconomic challenges to participating in a parenting program online. We found interesting differences in online component use based on cohort and incentives to participate online, but we did not randomize families to receive incentives (or not) at the beginning of the study. Incentives were added only in 2 later phases of the study. Randomized controlled trials examining the influence of incentives to participate in online supplements would provide important, additional evidence regarding their influence on participation. Additionally, in this study, the number of clicks was used as a proxy for completion of online modules; however, better measures of the time participants spend online and quality of those interactions are needed.

This study provides a foundation for future research on engaging parents in various contexts. While important differences were found when analyzed by parents' demographic characteristics, program incentives, and the ecological context, future research that includes parents' psychological characteristics, such as

motivation for participating in parenting programs, self-efficacy as a parent, and other unique individual characteristics, may further elucidate differences in participation. Additionally, research is needed to understand how engagement with online components versus face-to-face participation translates into behavior change (eg, improved parenting and child adjustment). Dismantling trials (which disaggregate different online components and face-to-face interventions options) as well as randomized preference trials (which randomize participants to choose or be assigned their modalities of choice) would provide practitioners and researchers with important information about what works in engaging and serving families through preventive parenting programs.

Conclusions

This research underscores the need to understand patterns of parent engagement in both Internet-based and face-to-face parenting programs. Our findings add to existing evidence that Internet-based participation has the potential to increase some parents' participation; however, face-to-face programming also appears to provide unique benefits to some parents, which could include instrumental and emotional support. Collecting data and monitoring the use of online components and face-to-face participation in parenting programs provides opportunities to better understand audiences and potentially improve both engagement and retention [7].

Acknowledgments

Authors' Contributions

JD and JR were the primary authors and examined the data reported in this study. SH and KP contributed to the writing, data management, and implementation of data for this paper. AG is the principal investigator on this study and provided ongoing supervision on the design, writing, and feedback for this paper.

Conflicts of Interest

AG is a consultant to Implementation Sciences International, a company that has a subcontract in this research.

Multimedia Appendix 1

Examples of web pages from ADAPT's online supplement.

[PPTX File, 974KB - jmir_v18i6e169_app1.pptx]

References

1. Lester P, Saltzman WR, Woodward K, Glover D, Leskin GA, Bursch B, et al. Evaluation of a family-centered prevention intervention for military children and families facing wartime deployments. Am J Public Health 2012 Mar;102 Suppl 1:S48-S54. [doi: 10.2105/AJPH.2010.300088] [Medline: 22033756]
2. Sandler IN, Schoenfelder EN, Wolchik SA, MacKinnon DP. Long-term impact of prevention programs to promote effective parenting: lasting effects but uncertain processes. Annu Rev Psychol 2011;62:299-329 [FREE Full text] [doi: 10.1146/annurev.psych.121208.131619] [Medline: 20822438]
3. Patterson GR, Forgatch MS, Degarmo DS. Cascading effects following intervention. Dev Psychopathol 2010 Nov;22(4):949-970 [FREE Full text] [doi: 10.1017/S0954579410000568] [Medline: 20883592]
4. Heinrichs N, Bertram H, Kuschel A, Hahlweg K. Parent recruitment and retention in a universal prevention program for child behavior and emotional problems: barriers to research and program participation. Prev Sci 2005 Dec;6(4):275-286. [doi: 10.1007/s11121-005-0006-1] [Medline: 16075192]
5. Thornton S, Calam R. Predicting intention to attend and actual attendance at a universal parent-training programme: a comparison of social cognition models. Clin Child Psychol Psychiatry 2011 Jul;16(3):365-383. [doi: 10.1177/1359104510366278] [Medline: 20876286]
6. Kazdin AE, Blase SL. Rebooting psychotherapy research and practice to reduce the burden of mental illness. Perspect Psychol Sci 2011 Feb 03;6(1):21-37 [FREE Full text] [doi: 10.1177/1745691610393527]
7. Cunningham CE, Deal K, Rimas H, Buchanan DH, Gold M, Sdao-Jarvie K, et al. Modeling the information preferences of parents of children with mental health problems: a discrete choice conjoint experiment. J Abnorm Child Psychol 2008 Oct;36(7):1123-1138. [doi: 10.1007/s10802-008-9238-4] [Medline: 18481167]
8. Metzler CW, Sanders MR, Rusby JC, Crowley RN. Using consumer preference information to increase the reach and impact of media-based parenting interventions in a public health approach to parenting support. Behav Ther 2012 Jun;43(2):257-270 [FREE Full text] [doi: 10.1016/j.beth.2011.05.004] [Medline: 22440064]

XSL•FO

RenderX

9. Breitenstein SM, Gross D, Christophersen R. Digital delivery methods of parenting training interventions: a systematic review. Worldviews Evid Based Nurs 2014 Jun;11(3):168-176. [doi: 10.1111/wvn.12040] [Medline: 24842341]

10. Kaplan K, Solomon P, Salzer MS, Brusilovskiy E. Assessing an Internet-based parenting intervention for mothers with a serious mental illness: a randomized controlled trial. Psychiatr Rehabil J 2014 Sep;37(3):222-231. [doi: 10.1037/prj0000080] [Medline: 24978623]

11. Wilson DK, Alia KA, Kitzman-Ulrich H, Resnicow K. A pilot study of the effects of a tailored web-based intervention on promoting fruit and vegetable intake in African American families. Child Obes 2014 Feb;10(1):77-84 [FREE Full text] [doi: 10.1089/chi.2013.0070] [Medline: 24299118]

12. Sanders MR, Baker S, Turner Karen M T. A randomized controlled trial evaluating the efficacy of Triple P Online with parents of children with early-onset conduct problems. Behav Res Ther 2012 Nov;50(11):675-684. [doi: 10.1016/j.brat.2012.07.004] [Medline: 22982082]

13. Enebrink P, Högström J, Forster M, Ghaderi A. Internet-based parent management training: a randomized controlled study. Behav Res Ther 2012 Apr;50(4):240-249. [doi: 10.1016/j.brat.2012.01.006] [Medline: 22398153]

14. Baggett KM, Davis B, Feil EG, Sheeber LB, Landry SH, Carta JJ, et al. Technologies for expanding the reach of evidence-based interventions: Preliminary results for promoting social-emotional development in early childhood. Topics Early Child Spec Educ 2010 Feb 1;29(4):226-238 [FREE Full text] [doi: 10.1177/0271121409354782] [Medline: 20454545]

15. Bert SC, Farris JR, Borkowski JG. Parent training: implementation strategies for adventures in parenting. J Prim Prev 2008 May;29(3):243-261. [doi: 10.1007/s10935-008-0135-y] [Medline: 18446440]

16. Scholer SJ, Hudnut-Beumler J, Dietrich MS. Why parents value a brief required primary care intervention that teaches discipline strategies. Clin Pediatr (Phila) 2012 Jun;51(6):538-545. [doi: 10.1177/0009922812439241] [Medline: 22496174]

17. Taylor TK, Webster-Stratton C, Feil EG, Broadbent B, Widdop CS, Severson HH. Computer-based intervention with coaching: an example using the Incredible Years program. Cogn Behav Ther 2008;37(4):233-246 [FREE Full text] [doi: 10.1080/16506070802364511] [Medline: 18803072]

18. Sanders M, Calam R, Durand M, Liversidge T, Carmont SA. Does self-directed and web-based support for parents enhance the effects of viewing a reality television series based on the Triple P-Positive Parenting Programme? J Child Psychol Psychiatry 2008 Sep;49(9):924-932. [doi: 10.1111/j.1469-7610.2008.01901.x] [Medline: 18492040]

19. Calam R, Sanders MR, Miller C, Sadhnani V, Carmont S. Can technology and the media help reduce dysfunctional parenting and increase engagement with preventative parenting interventions? Child Maltreat 2008 Nov;13(4):347-361. [doi: 10.1177/1077559508321272] [Medline: 18641169]

20. Brouwer W, Kroeze W, Crutzen R, de NJ, de Vries Nanne K, 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 Jan;13(1):e2 [FREE Full text] [doi: 10.2196/jmir.1639] [Medline: 21212045]

21. Doty JL, Dworkin J, Connell JH. Examining digital differences: Parents' online activities. Family Science Review 2012 Dec;17(2):18-39 [FREE Full text]

22. Radey M, Randolph KA. Parenting sources: How do parents differ in their efforts to learn about parenting? Fam Relat 2009 Dec;58(5):536-548 [FREE Full text] [doi: 10.1111/j.1741-3729.2009.00573.x]

23. Rothbaum F, Martland N, Jannsen JB. Parents' reliance on the Web to find information about children and families: socio-economic differences in use, skills and satisfaction. J Appl Dev Psychol 2008 Mar;29(2):118-128 [FREE Full text] [doi: 10.1016/j.appdev.2007.12.002]

24. Cunningham CE, Chen Y, Deal K, Rimas H, McGrath P, Reid G, et al. The interim service preferences of parents waiting for children's mental health treatment: a discrete choice conjoint experiment. J Abnorm Child Psychol 2013 Feb;41(6):865-877. [doi: 10.1007/s10802-008-9238-4]

25. Kennedy G, Judd T, Dalgarno B, Waycott J. Beyond natives and immigrants: exploring types of net generation students. J Comput Assist Lear 2010 Oct;26(5):332-343 [FREE Full text] [doi: 10.1111/j.1365-2729.2010.00371.x]

26. Nasah A, DaCosta B, Kinsell C, Seok S. The digital literacy debate: an investigation of digital propensity and information and communication technology. Educ Technol Res Dev 2010 Oct;58(5):531-555 [FREE Full text] [doi: 10.1007/s11423-010-9151-8]

27. Brady E, Guerin S. "Not the romantic, all happy, coochy coo experience": a qualitative analysis of interactions on an Irish parenting web site. Fam Relat 2010 Feb;59(1):14-27 [FREE Full text] [doi: 10.1111/j.1741-3729.2009.00582.x]

28. Sarkadi A, Bremberg S. Socially unbiased parenting support on the Internet: a cross-sectional study of users of a large Swedish parenting website. Child Care Health Dev 2005 Jan;31(1):43-52. [doi: 10.1111/j.1365-2214.2005.00475.x] [Medline: 15658965]

29. Creasey RR, Trikha S. Research Development and Statistics Directorate Home Office Online Report 48/04. 2004. Meeting parents' needs for information: evidence from the 2001 Home Office Citizenship Survey URL: http://webarchive. nationalarchives.gov.uk/20120919132719/http://www.communities.gov.uk/documents/communities/pdf/452422.pdf [accessed 2014-11-14] [WebCite Cache ID 6blP8o9St]

30. McBride BA, Brown GL, Bost KK, Shin N, Vaughn B, Korth B. Paternal identity, maternal gatekeeping and father involvement. Fam Relat 2005 Jul;54(3):360-372 [FREE Full text]

31. Crouter AC, McHale S. Temporal rhythms in family life: seasonal variation in the relation between parental work and family processes. Dev Psychol 1993 Mar;29(2):198-205 [FREE Full text] [doi: 10.1037/0012-1649.29.2.198.x]

32. Baker CN, Arnold DH, Meagher S. Enrollment and attendance in a parent training prevention program for conduct problems. Prev Sci 2011 Jun;12(2):126-138. [doi: 10.1007/s11121-010-0187-0] [Medline: 21052834]

33. Coatsworth JD, Duncan LG, Pantin H, Szapocznik J. Patterns of retention in a preventive intervention with ethnic minority families. J Prim Prev 2006 Mar;27(2):171-193 [FREE Full text] [doi: 10.1007/s10935-005-0028-2] [Medline: 16532263]

34. Reyno SM, McGrath PJ. Predictors of parent training efficacy for child externalizing behavior problems--a meta-analytic review. J Child Psychol Psychiatry 2006 Jan;47(1):99-111. [doi: 10.1111/j.1469-7610.2005.01544.x] [Medline: 16405646]

35. Peters S, Calam R, Harrington R. Maternal attributions and expressed emotion as predictors of attendance at parent management training. J Child Psychol Psychiatry 2005 Apr;46(4):436-448. [doi: 10.1111/j.1469-7610.2004.00365.x] [Medline: 15819652]

36. Gross D, Julion W, Fogg L. What motivates participation and dropout among low-income urban families of color in a prevention intervention? Fam Relat 2001 Jul;50(3):246-254 [FREE Full text] [doi: 10.1111/j.1741-3729.2001.00246.x]

37. Spoth R, Redmond C, Shin C. Modeling factors influencing enrollment in family-focused preventive intervention research. Prev Sci 2000 Dec;1(4):213-225. [Medline: 11523749]

38. Kazdin AE, Holland L, Crowley M. Family experience of barriers to treatment and premature termination from child therapy. J Consult Clin Psychol 1997 Jun;65(3):453-463. [Medline: 9170769]

39. Nix RL, Bierman KL, McMahon RJ, Conduct Problems Prevention Research Group. How attendance and quality of participation affect treatment response to parent management training. J Consult Clin Psychol 2009 Jun;77(3):429-438 [FREE Full text] [doi: 10.1037/a0015028] [Medline: 19485585]

40. Coatsworth JD, Duncan LG, Pantin H, Szapocznik J. Retaining ethnic minority parents in a preventive intervention: the quality of group process. J Prim Prev 2006 Jul;27(4):367-389. [doi: 10.1007/s10935-006-0043-y] [Medline: 16802072]

41. Fox DP, Gottfredson DC. Differentiating completers from non-completers of a family-based prevention program. J Prim Prev 2003 Dec;24(2):111-124 [FREE Full text] [doi: 10.1023/A:1025988227486]

42. Beatty SE, Cross DS. Investigating parental preferences regarding the development and implementation of a parent-directed drug-related educational intervention: an exploratory study. Drug Alcohol Rev 2006 Jul;25(4):333-342. [doi: 10.1080/09595230600741172] [Medline: 16854659]

43. Al-Halabí Díaz S, Errasti Pérez JM. Use of small incentives for increasing participation and reducing dropout in a family drug-use prevention program in a Spanish sample. Subst Use Misuse 2009 Dec;44(14):1990-2000. [doi: 10.3109/10826080902844870] [Medline: 20001690]

44. Axford N, Lehtonen M, Kaoukji D, Tobin K, Berry V. Engaging parents in parenting programs: lessons from research and practice. Children Youth Serv Rev 2012 Oct;34(10):2061-2071 [FREE Full text] [doi: 10.1016/j.childyouth.2012.06.011]

45. Heinrichs N. The effects of two different incentives on recruitment rates of families into a prevention program. J Prim Prev 2006 Jul;27(4):345-365. [doi: 10.1007/s10935-006-0038-8] [Medline: 16802074]

46. Dumas JE, Begle AM, French B, Pearl A. Effects of monetary incentives on engagement in the PACE parenting program. J Clin Child Adolesc Psychol 2010;39(3):302-313 [FREE Full text] [doi: 10.1080/15374411003691792] [Medline: 20419572]

47. Snow JN, Frey MR, Kern RM. Attrition, financial incentives, and parent education. The Fam J 2002 Oct;10(4):373-378 [FREE Full text] [doi: 10.1177/106648002236755]

48. Bubolz MM, Sontag MS. Human ecology theory. In: Boss PG, Doherty WJ, LaRossa R, Schumm WR, Steinmetz SK, editors. Sourcebook of Family Theories and Methods: A Contextual Approach. New York: Springer; 1993.

49. McCurdy K, Daro D. Parent involvement in family support programs: an integrated theory. Fam Relat 2001 Apr;50(2):113-121 [FREE Full text] [doi: 10.1111/j.1741-3729.2001.00113.x]

50. Bowen GL, Mancini JA, Martin JA, Ware WB, Nelson JP. Promoting the adaptation of military families: an empirical test of a community practice model. Fam Relat 2003 Jan;52(1):33-44 [FREE Full text] [doi: 10.1111/j.1741-3729.2003.00033.x]

51. Zickuhr K, Smith A. Pew Research Center. 2012 Apr. Digital Differences URL: http://www.pewinternet.org/2012/04/13/digital-differences/ [accessed 2012-05-02] [WebCite Cache ID 6X7Its9cq]

52. Gottman JM, Gottman JS, Atkins CL. The Comprehensive Soldier Fitness program: family skills component. Am Psychol 2011 Jan;66(1):52-57. [doi: 10.1037/a0021706] [Medline: 21219048]

53. Schachman KA. Online fathering: the experience of first-time fatherhood in combat-deployed troops. Nurs Res 2010 Feb;59(1):11-17. [doi: 10.1097/NNR.0b013e3181c3ba1d] [Medline: 20010040]

54. Gewirtz AH, Pinna KLM, Hanson SK, Brockberg D. Promoting parenting to support reintegrating military families: after deployment, adaptive parenting tools. Psychol Serv 2014 Feb;11(1):31-40 [FREE Full text] [doi: 10.1037/a0034134] [Medline: 24564441]

55. Samper RE, Taft CT, King DW, King LA. Posttraumatic stress disorder symptoms and parenting satisfaction among a national sample of male Vietnam veterans. J Trauma Stress 2004 Aug;17(4):311-315. [doi: 10.1023/B:JOTS.0000038479.30903.ed] [Medline: 15462538]

56. Forgatch MS, DeGarmo DS. Parenting through change: an effective prevention program for single mothers. J Consult Clin Psychol 1999 Oct;67(5):711-724. [Medline: 10535238]

57. Gewirtz AH, Erbes CR, Polusny MA, Forgatch MS, Degarmo DS. Helping military families through the deployment process: strategies to support parenting. Prof Psychol Res Pr 2011 Feb;42(1):56-62 [FREE Full text] [doi: 10.1037/a0022345] [Medline: 21841889]

58. StataCorp. 2015. Stata Base Reference Manual: Release 14 URL: http://www.stata.com/manuals14/r.pdf [accessed 2015-09-28] [WebCite Cache ID 6buQXZF9i]

Abbreviations

ADAPT: After Deployment, Adaptive Parenting Tools
NGR: National Guard and Reserves
PMTO: Parent Management Training–Oregon Model
RCT: randomized controlled trial
SES: socioeconomic status

XSL•FO
RenderX

<u>Original Paper</u>

Decreased Body Mass Index in Schoolchildren After Yearlong Information Sessions With Parents Reinforced With Web and Mobile Phone Resources: Community Trial

Jenny Vilchis-Gil[1], PhD (c); Miguel Klünder-Klünder[1], PhD; Ximena Duque[2], PhD; Samuel Flores-Huerta[1], MD

[1]Hospital Infantil de México Federico Gómez, Ministry of Health (SSA), Community Health Research Department, Mexico City, Mexico

[2]Mexican Institute of Social Security, Unit of Medical Research in Infectious and Parasitic Diseases, Mexico City, Mexico

Corresponding Author:
Samuel Flores-Huerta, MD
Hospital Infantil de México Federico Gómez, Ministry of Health (SSA)
Community Health Research Department
Dr. Márquez No. 162
Mexico City, 06720
Mexico
Phone: 52 55 5228 9917 ext 4510
Fax: 52 55 5228 9917
Email: floreshuertamd@gmail.com

Abstract

Background: The obesity pandemic has now reached children, and households should change their lifestyles to prevent it.

Objective: The objective was to assess the effect of a comprehensive intervention on body mass index z-score (BMIZ) in schoolchildren.

Methods: A yearlong study was conducted at 4 elementary schools in Mexico City. Intervention group (IG) and control group (CG) were split equally between governmental and private schools. Three educational in-person parents and children sessions were held at 2-month intervals to promote healthy eating habits and exercise. To reinforce the information, a website provided extensive discussion on a new topic every 2 weeks, including school snack menus and tools to calculate body mass index in children and adults. Text messages were sent to parents' mobile phones reinforcing the information provided. The IG contained 226 children and CG 181 children. We measured their weight and height and calculated BMIZ at 0, 6, and 12 months.

Results: The CG children showed a change of +0.06 (95% CI 0.01, 0.11) and +0.05 (95% CI 0.01, 0.10) in their BMIZ at 6 and 12 months, respectively. The BMIZ of IG children decreased by -0.13 (95% CI -0.19 to -0.06) and -0.10 (95% CI -0.16 to -0.03), respectively, and the effect was greater in children with obesity.

Conclusions: The comprehensive intervention tested had beneficial effects, preserved the BMIZ of normal weight children, and reduced the BMIZ of children with obesity.

Introduction

Childhood obesity is a public health problem because it is associated with comorbidities such as metabolic syndrome, hypertriglyceridemia, and type 2 diabetes [1-3]. Therefore, it is important to develop alternative strategies to prevent this health problem from an early age. In the last two decades, overweight and obesity have increased in Mexican children aged 5-11 years. The national nutrition surveys showed that from 1999 to 2006

the prevalence increased from 26.9% to 34.8%. In 2012, the prevalence was 34.4% [4], indicating that this problem prevails with unacceptably high figures for this age group [5].

Metabolically, obesity is caused by an imbalance in which energy intake exceeds expenditure in a persistent fashion. Low energy expenditure is caused by decreased physical activity and increased sedentary habits [6-8]. To neutralize this problem and its comorbidities, intervention studies have been performed to change lifestyles of eating and physical activity, but many have

XSL•FO
RenderX

been conducted in minority populations [9] without a control group [10,11] for short periods [9,12] and have had no lasting effect on the body mass index z-scores (BMIZ) [13,14]. The energy imbalance of overweight children results from the unhealthy lifestyles acquired at home, and health providers need to inform parents on how to change unhealthy feeding habits and sedentary habits for better alternatives.

Studies in children and adolescents have shown that Web-based information improves people's knowledge, attitudes, and behaviors related to feeding and physical activity [15,16]. However, such information is rarely a comprehensive educational intervention [10] that also incorporates sociocultural and economic conditions of the target group. Furthermore, there is a need to educate parents about their role in the prevention of obesity. It is expected that information would change feeding behaviors, the types of foods available at home, and preparation and enjoyment of healthy menus. Currently, the Internet and mobile phones are regular informational tools of many households in Mexico City with several apps available for users. In Mexico, the percentage of Internet users nationwide grew from 19.5% to 40.2% from 2006 to 2012, the largest number being in urban areas. In the same period, the rate of mobile phone users increased from 52.6% to 90.8% [17]. One advantage of mobile phones for educational purposes is the instantaneous relay of short messages at any time and place.

The aim of this study was to compare the daily nutrition-based activities performed in schools in Mexico City, such as the availability of healthy food and drinking water [18,19], with an educational intervention involving in-person sessions with parents, distance activities using Web resources and mobile phone reminders, and printed materials in order to improve or at least preserve the BMIZ of Mexican children in elementary schools.

Methods

Design and Study Population

The study was approved by Research, Ethics, and Biosafety Committees of the Federico Gomez Children's Hospital of Mexico (HIMFG in Spanish), a National Institute of Health. After the Ministry of Public Education granted the authorization to perform the study, 4 elementary schools from a middle-class suburb of Mexico City with similar number of students were included in the study. All school principals granted authorization for the study. The intervention was implemented in 2 schools, 1 governmental and 1 private (intervention group, IG), whereas

the other 2 schools, 1 governmental and 1 private, were controls (control group, CG). Before implementation, educational materials and the website were designed and developed, which were the tools of the intervention. The subsequent activities are described as phase 1 of the study. The grades that participated were the first to fourth. Boys and girls were included, regardless of their BMIZ. Before starting the study, both objectives and activities were explained to teachers, parents, and children and also verbal agreement and written informed consent was obtained from children and their parents, respectively.

Phase 1: Design and Development of Educational Materials and Website

Artwork That Would Accompany the Messages

The artworks for this project, such as images of children eating healthy foods, were created by the designers of Universum, the Science Museum of the National Autonomous University of Mexico (UNAM). Designs were considered to be ad hoc for the age of the children and culture of Mexico City.

Informative Topics, Posters, and Messages

Twenty topics were developed to inform parents about overweight, obesity, healthy eating, physical activity, and health risks that obesity involves, which are described in Table 1 [20-25]. The topics were developed by pediatricians, nutritionists, nurses, and physical educators, who were asked to adjust the information to the following format: not more than 500 words, to suggest activities that could be implemented by the family, and annotate 2 or more hyperlinks to obtain more information, if desired. Topics were edited by a pediatric nutrition expert, for the purpose of standardizing the language and isolating 2 brief messages that would be sent to the parents' mobile phones. Information posters were developed to reinforce the written topics, which were placed in strategic locations within the intervention schools.

Various materials were developed for the children to take home, such as laminated place mats with the images of the Mexican Eatwell Plate [26] and Drinkwell Pitcher [27] and the pyramid of physical activity. For illustrative purposes, some screenshots of the educational materials used in the intervention are presented in Multimedia Appendix 1 (Selected screenshots of the materials used in the intervention); additionally, the authors can be contacted to obtain samples of the materials.

Guidelines for parents were developed with information on how to prepare a healthy school lunch, including numerous examples.

Figure 1. Study population: screening results and follow-up.
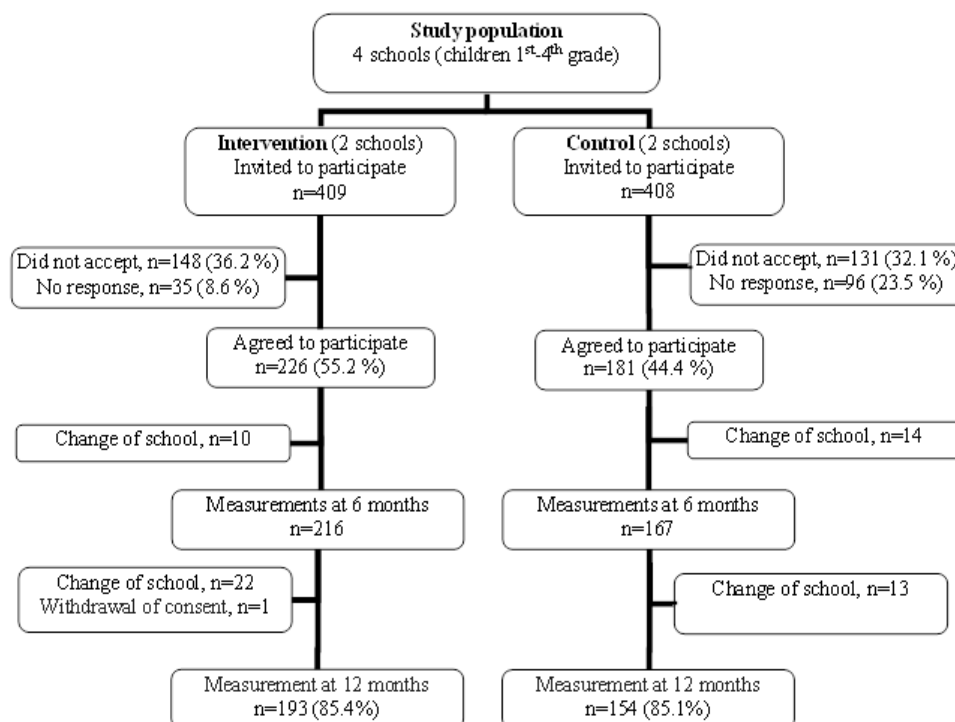
Table 1. Work topics on the website, phone messages, and posters.

Website topics	Messages	Poster
1. What is obesity? Why are there more obese people now than before?	M1, M2	The change in lifestyle has affected our health
2. Obesity and its relationship to chronic degenerative diseases	M3, M4	
3. Lifestyles, feeding, activity/sedentary, and health monitoring with regular measurements	M5, M6	Consequences of overweight and obesity in health
4. Lifestyle and health. We stay as we are or we make an effort to improve the health of the family and of each of its members	M7, M8	
5. Eating breakfast and bringing lunch to school are essential for health and school performance of children	M9, M10	Breakfast at home and lunch at school
6. Physical activity is much more than expending energy	M11, M12	
7. Sedentary lifestyle, a risk for health	M13, M14	Free and spontaneous play
8. Measure yourself. Self-measurement is much more than knowing your weight	M15, M16	
9. Plain water is the healthiest	M17, M18	How can we avoid a sedentary lifestyle?
10. Not all the energy that we consume is the same in terms of health	M19, M20	
11. Carbohydrates. Health benefits and risks	M21, M22	Natural water is the healthiest drink
12. Fiber. Health benefits and risks	M23, M24	
13. Importance of eating fruits and vegetables	M25, M26	Fruits and vegetables in my feeding
14. Lipids. Health benefits and risks	M27, M28	
15. Proteins in food. Health benefits and risks	M29, M30	Varied feeding
16. Table salt in food. Health benefits and risks	M31, M32	
17. Vitamins and minerals. Health benefits and risks	M33, M34	Natural foods better than processed foods
18. Family behavior during food consumption at home	M35, M36	
19. Planning food purchases. Learning to read labels on processed and industrialized foods	M37, M38	
20. Integration. Create a healthy eating environment with your family	M39, M40	

Design and Development of the Website

The project website was nested in the official HIMFG website www.himfg.edu.mx. The picture of 2 children eating healthy foods was the identification icon of the project. Upon accessing the site, the user found the biweekly highlighted topic and the following icons:

1. Previous topics. By accessing this icon, the list of biweekly topics appeared, which could be accessed at any time. At the end of 2 weeks, the current theme was stored in this folder.

2. School snacks. This icon contained information on how to develop a healthy school snack and 25 examples of school snacks.

3. Guidelines on healthy eating and physical activity. This icon contained files of 2 guides on nutrition and physical activity; these printed guides were also sent to parents through their children.

4. Posters. This folder contained the posters hung at schools on a monthly basis. Each poster reinforced the current theme. A total of 9 posters were placed.

5. Software for the calculation and interpretation of body mass index (BMI). This icon invited parents to assess the BMI of themselves and their children. When they entered a weight and height, the program returned a BMI, indicating whether there was a health risk.

6. Contact, questions, and comments. In this section, parents could contact researchers to raise questions or make comments related to the project.

7. Window for researchers. The page included a window that allowed researchers to activate the user, evaluate the number of accesses to the site by each of the families, read the comments or questions that were raised by parents, and keep the page updated with the information of each of the icons. A visual of the website is shown in Multimedia Appendix 2 (Selected screenshot of the website).

Phase 2: Intervention Implementation

The intervention was implemented from October 2013 to July 2014; parents and all children, regardless of their BMIZ, participated in the intervention.

Activities of Parents in the Intervention Group

1. Website. An interactive workshop to learn and use the project website was developed. The name of their child was the username, and the password was the date of birth of their child. Parents were invited to access the site at least once every 2 weeks and review the current topic.

2. Messages by mobile phone. A short message (25 words on average) was sent to the parents' mobile phone, every week. The message was related to the current topic of the website that motivated and reinforced behavior changes. A total of 40 messages were sent (Multimedia Appendix 3: Selected screenshots of examples of messages sent to parents' mobile phone).

3. In-person activities. Parents from the intervention schools were invited to 3 sessions of 1 hour each, once every 2 months, for the purpose of giving them feedback on topics of eating and physical activity that were on the website, to participate in some project activities, and to answer their questions and receive comments to modify or improve the website.

Activities of Children in the Intervention Group

1. Workshops. Two nutritionists and a physical educator conducted 4 workshops, 1 every 2 months, with a duration of 1.5 hours. The workshops were integrated with both board games and physical games and with educational materials to reinforce healthy eating habits and physical activity.

2. Educational materials. They were provided with board games and plastic place mats with pictures of a healthy dish [26], the pyramid of physical activity, and a picture of a healthy drink pitcher [27].

3. Visit to Universum Museum. Children and parents visited the Life in Balance room of Universum, the Science Museum of the UNAM.

4. Posters. Each month, in a visible area and an area with large influx of children inside the school (eg, at the entrance of the school and in the schoolyard) a new poster alluding to the project and specifically the current website topic was placed.

Measurements

Sociodemographic Information

Information about the socioeconomic status was obtained by administering a questionnaire to know the number of children, parental education, and housing characteristics.

Anthropometric Measurements

Weight, height, and waist circumference were measured in schoolchildren of all schools at baseline and then at 6 and 12 months. Weight was measured with a digital scale (Seca 882; Seca Corp, Hamburg, Germany) with an accuracy of 0.1 kg. Height was measured with a stadiometer (Seca 225; Seca Corp, Hamburg, Germany) with an accuracy of 0.1 cm. Waist circumference was measured with a nonelastic flexible measuring tape (Seca 200). Two trained nutritionists took these measurements according to international procedures [28]. Briefly, the children were measured without shoes and in light clothing, standing in the center of the scale platform or stadiometer, arms resting loosely on the sides, and head positioned in the Frankfort plane. To measure waist circumference, the child climbed an anthropometric box 60 cm tall, the tape was placed at the midpoint between the iliac crest and the lower costal margin, and the reading was taken at the end of a normal exhalation. The purpose of having the child stand on the box was to bring the reader's eyes close to the height of the tape to avoid parallax errors.

Body mass index (BMI) and BMI z-score (BMIZ) are measures used to define obesity. BMI is a measure of weight adjusted for height. It is calculated as weight in kilograms divided by the square of height in meters. Whereas in adults the BMI cut points that define obesity and overweight are not linked to age and do not differ for males and females, in growing children BMI varies

with age and sex. For BMI to be meaningful in children it must be compared to a reference-standard that accounts for child age and sex. Body mass index z-scores, also called BMI standard deviation (SD) scores, are measures of relative weight adjusted for child age and sex. Given a child's age, sex, BMI, and an appropriate reference standard, a BMI z-score (or its equivalent BMI-for-age percentile) can be determined.

Children were classified according to the nutritional status as underweight (z-score ≤ -2), normal weight (-2< z-score <1), overweight (1≤ z-score <2), or obese (z-score ≥2), using as a reference the data from the World Health Organization report of 2007 [29]. Additionally, the waist circumference percentile was calculated with consideration of age, sex, and height, using the tables for Mexican children [30].

After each of the anthropometric measurements, both groups of children were handed a letter with the results of the nutritional status and tips to maintain or improve their health. The recommendations mainly focused on eating and physical activity habits.

Data Analyses

Descriptive statistics were used to describe the baseline of the study population. The mean weight and height were adjusted for age and sex by multiple linear regression. The socioeconomic variable was built with the information obtained from the questionnaire on housing characteristics and ownership of property. Households were grouped into tertiles according to their score for socioeconomic status. To compare the groups at baseline, t test for independent data was used for continuous variables and the χ^2 test was used for categorical data.

Because, frequently in this type of study, participants do not always follow instructions and consequently adherence is never 100%, we considered it appropriate to perform the data analysis by intention-to-treat, in which all participants assigned to each group are analyzed regardless of their adherence.

The intragroup BMIZ changes from baseline to 6 months and from baseline to 12 months were compared using paired t test. To compare the intergroup BMIZ change, t test for independent data was used.

A model of mixed-effects linear regression was used to assess the change in BMIZ during follow-up. Because children are nested within schools, a mixed model with 3 levels and random intercepts by subject and school were tested. The model was adjusted by the fixed variables of age and BMIZ at baseline. The interaction between the intervention group and the time at each evaluation, 6 and 12 months, was assessed. Mean BMIZ by group (intervention or control) and by time was calculated and a graphic was done using marginal analysis. P values < .05 were considered statistically significant for all analyses. The analysis was performed using Stata v12.0 (StataCorp, College Station, TX, USA).

Results

The number of schoolchildren who participated was 226 in the IG and 181 in the CG. The participation rates were 55.2% (226/409) in IG and 44.4% (181/408) in CG. At 12 months, 85% of children in both groups completed the measurements. The proportion of children who did not conclude the study can be attributed to children moving to different schools for personal reasons. However, there was no movement of children between study groups (Figure 1).

The development of the intervention is presented in Table 2; among the parents, 51.3% (116/226) attended at least one educational session, 40.7% (92/226) consulted the website, and 91.2% (206/226) received mobile phone messages. All children handed in materials and participated in workshops.

Table 2. Intervention development during the school year (n=226).

Intervention			n (%)
Parents' sessions			
	None		110 (48.7)
	At least one session		116 (51.3)
		1 session	78 (34.5)
		2 sessions	34 (15.0)
		3 sessions	4 (1.8)
Consultation of the website			
	No		134 (59.3)
	Yes		92 (40.7)
Sending messages to mobile phone			
	No message		20 (8.8)
	Message		206 (91.2)

Table 3 presents the baseline characteristics of the study population. The demographic, anthropometric, and socioeconomic characteristics of both groups were similar. The average age was 8 years, and there was no difference in BMIZ. The only exception was the use of mobile phones, which was significantly higher in the IG.

Table 3. Baseline characteristics of the study population according to study group.

Characteristics	Control n=181	Intervention n=226	P [a]
Age in years, mean (SD)	8.1 (1.2)	7.9 (1.2)	.13
Sex, female, n (%)	90 (49.7)	101 (44.7)	.31
Anthropometric measurements			
Weight[b], kg, mean (SD)	30.1 (5.1)	29.3 (5.0)	.13
Height[b], cm, mean (SD)	127.4 (7.5)	126.4 (7.4)	.14
BMI[c] z-score[d], mean (SD)	0.98 (1.3)	0.85 (1.4)	.35
BMI classification[d], n (%)			
Normal (-2≤ z-score <1)	89 (49.2)	129 (57.1)	
Overweight (1≤ z-score <2)	46 (25.4)	52 (23.0)	
Obese (z-score ≥2)	45 (25.4)	45 (19.9)	.25
Waist circumference, percentile, mean (SD)	53.0 (21.8)	52.2 (21.2)	.69
Maternal schooling, n (%)			
Secondary or less	24 (14.1)	42 (20.8)	
High school or technical school	77 (45.3)	76 (37.6)	
College career or postgraduate	69 (40.6)	84 (41.6)	.16
Number of children, n (%)			
1-2	116 (72.1)	153 (76.1)	
≥3	45 (27.9)	48 (23.9)	.38
Socioeconomic status (tertiles), n (%)			
Lower	42 (26.1)	63 (32.1)	
Medium	53 (32.9)	76 (38.8)	
Higher	66 (41.0)	57 (29.1)	.06
Availability of Internet, n (%)			
At home	129 (86.6)	141 (82.9)	.37
On the mobile phone	94 (63.0)	92 (54.1)	.11
With mobile phone, n (%)			
Mothers	146 (85.4)	190 (92.2)	.03
Parents	155 (90.1)	201 (95.3)	.05

[a] t test and χ^2 test.

[b] Means adjusted by age and sex.

[c] BMI: body mass index.

[d] World Health Organization, 2007.

Table 4 shows the differences in BMI and BMIZ both within and between groups from baseline to 6 months and from baseline to 12 months. Intragroup analysis showed that in the IG, BMIZ decreased from 0 to 6 months (-0.07, 95% CI -0.12 to -0.03) and was maintained at 12 months. Whereas BMIZ of the CG increased (0.07, 95% CI 0.02, 0.12) from baseline to 6 months and persisted until 12 months. When stratified by nutritional status according to baseline BMIZ, the BMIZ of CG children with normal weight increased at 6 months (P=.002) and was maintained at 12 months (P=.04), whereas the BMIZ of the IG was maintained throughout the follow-up. Overweight children of both groups did not show significant changes from baseline to 6 months or from baseline to 12 months. Concerning obese children, those of CG maintained their BMIZ throughout the follow-up, whereas BMIZ of IG children decreased at 6 months (P<.001) and continued decreasing to 12 months (P=.001).

When comparing the BMIZ change between IG and CG in children who started the study with normal weight, differences observed between groups at 6 and 12 months were -0.17 (95% CI -0.27 to -0.07) and -0.12 (95% CI -0.24 to -0.01), respectively. In children who were overweight at baseline, an effect between groups of -0.15 (95% CI -0.24 to -0.01) at 6 months was observed, although this effect was not maintained at 12 months. In obese children, the effect on BMIZ between groups was -0.12 (95% CI -0.23 to -0.02) at 6 months and -0.16 (95% CI -0.32 to -0.01) at 12 months.

Table 4. Intra- and intergroup body mass index *z*-score comparison during follow-up (intervention group, n=191; control group, n=154).

Characteristics		0 months	6 months	12 months	Δ 0 to 6 months	*P* [a]	Δ 0 to 12 months	*P* [a]
		Mean (SD)	Mean (SD)	Mean (SD)	Mean (95% CI)		Mean (95% CI)	
BMI[b] (kg/m^2)								
	Control	18.4 (3.2)	18.8 (3.3)	19.1 (3.5)	0.50 (0.38 to 0.60)	<.001	0.77 (0.62 to 0.92)	<.001
	Intervention	17.9 (3.3)	18.0 (3.4)	18.4 (3.6)	0.13 (0.03 to 0.23)	.01	0.50 (0.36 to 0.63)	<.001
BMIZ[c,d]								
	Control	1.01 (1.3)	1.08 (1.3)	1.06 (1.3)	0.07 (0.02 to 0.12)	.004	0.05 (-0.01 to 0.11)	.07
	Intervention	0.85 (1.3)	0.77 (1.3)	0.80 (1.3)	-0.07 (-0.12 to -0.03)	.002	-0.05 (-0.10 to 0.01)	.08
BMIZ according to baseline nutritional status[d]								
Normal weight								
	Control	-0.15 (0.7)	-0.01 (0.8)	-0.05 (0.8)	0.13 (0.05 to 0.21)	.002	0.10 (0.01 to 0.19)	.04
	Intervention	-0.13 (0.7)	-0.17 (0.7)	-0.16 (0.8)	-0.04 (-0.10 to 0.02)	.20	-0.02 (-0.10 to 0.05)	.46
	Difference IG[e] versus CG[f]				-0.17 (-0.27 to -0.07)	<.001	-0.12 (-0.24 to -0.01)	.03
Overweight								
	Control	1.43 (0.3)	1.52 (0.4)	1.52 (0.5)	0.09 (-0.01 to 0.18)	.07	0.09 (-0.03 to 0.21)	.15
	Intervention	1.52 (0.3)	1.46 (0.4)	1.56 (0.4)	-0.06 (-0.16 to 0.03)	.20	0.05 (-0.06 to 0.15)	.36
	Difference IG versus CG				-0.15 (-0.24 to -0.01)	.03	-0.04 (-0.20 to 0.11)	.59
Obesity								
	Control	2.64 (0.5)	2.60 (0.5)	2.58 (0.5)	-0.05 (-0.11 to 0.02)	.14	-0.06 (-0.15 to 0.04)	.22
	Intervention	2.78 (0.6)	2.61 (0.7)	2.55 (0.7)	-0.17 (-0.26 to -0.08)	<.001	-0.22 (-0.35 to -0.09)	.001
	Difference IG versus CG				-0.12 (-0.23 to -0.02)	.02	-0.16 (-0.32 to -0.01)	.04

[a] Paired *t* test and *t* test for independent samples.

[b] BMI: body mass index.

[c] BMIZ: body mass index *z*-score.

[d] World Health Organization, 2007.

[e] IG: intervention group.

[f] CG: control group.

On the basis of a linear regression model with random intercept mixed effect and adjusted by baseline age and BMIZ, the graph shown in Figure 2 was constructed. The continuous line in this figure shows that in the first 6 months the intervention had a greater effect on the BMIZ of schoolchildren; whereas BMIZ increased in the CG children (dashed line). At baseline, IG and CG children began with a similar BMIZ; however, the mean of BMIZ changed at 6 and 12 months between groups. The adjusted means of BMIZ of the CG were estimated at 0.93 (95% CI 0.89-0.97), 0.99 (95% CI 0.95-1.04), and 0.98 (95% CI 0.94-1.03) at baseline, 6 months, and 12 months, respectively; those of IG were estimated at 0.92 (95% CI 0.89-0.96), 0.86 (95% CI 0.82-0.89), and 0.88 (95% CI 0.84-0.92), respectively.

Figure 2. Body mass index *z*-score (BMIZ) change in children during follow-up. Linear regression mixed effects model with random intercept, adjusted for baseline age and baseline BMIZ.
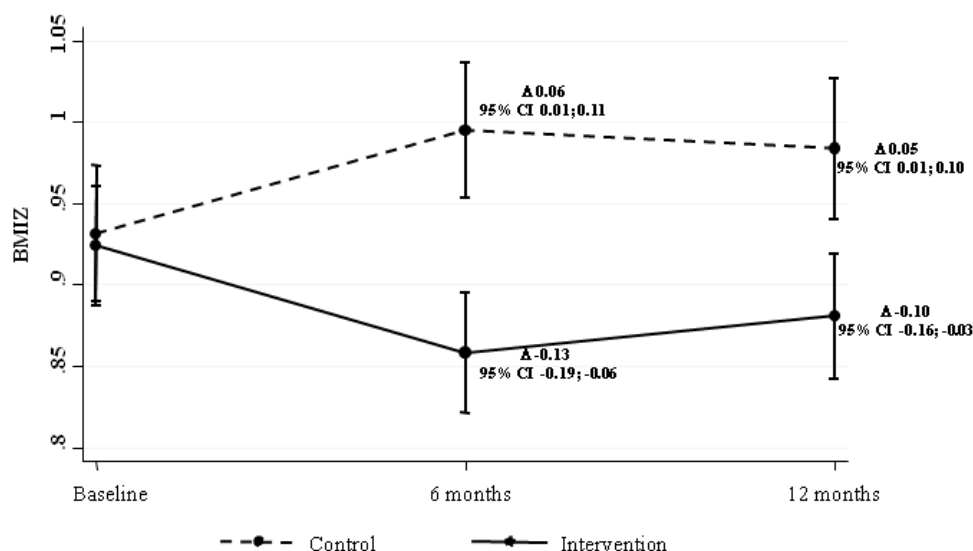
Discussion

This study shows the results of an intervention implemented in elementary schools comprising multiple components. The intervention included in-person and remote activities aimed at parents and children. Comparing the BMIZ during the study with the BMIZ at baseline, in the intervention group BMIZ decreased and in the control group, BMIZ increased. In addition, the intervention maintains the BMIZ of normal weight children and decreased the BMIZ of obese children. Moreover, it is pertinent to emphasize that the strategy described in this report was not intended to treat obesity but sought to reach as many parents as possible to promote changes in the eating and physical activity habits of their children regardless of BMIZ.

For the management of obese adolescents and adults, the effectiveness of sending information electronically versus individual face-to-face consultation has been compared, and weight reduction is greater in those receiving individual consultation [31]. However, such consultations become unsustainable because of the magnitude of the problem and also because each subject requires a large amount of attention and resources. In this study, the effect of the yearlong follow-up intervention in the whole sample was a BMIZ reduction of 0.10, equivalent to a decrease of 0.4 kg. It is important to mention that studies with similar interventions achieved beneficial effects on BMIZ, whereas others achieved positive changes only in eating and physical activity habits and not the BMIZ [17,32,33].

In this report, when the effect was analyzed by subgroups considering the initial BMIZ, the intervention had a differential effect: the BMIZ of IG children with normal weight was maintained during the follow-up, whereas BMIZ of children from the CG increased. The obese children from the IG reduced 0.22 of BMIZ by 12 months; this effect was greater than that observed in obese children of the CG (-0.06) and the whole population of the intervention group (-0.10). These results are similar to those of a Chilean intervention study that found that

obese schoolchildren are the subpopulation with better responses [34].

In our study, in the children with overweight in the intervention group the BMIZ decreased at 6 months (intervention effect) but this effect disappeared at 12 months; this may be because the parents of these children do not yet perceive the health problem in which their children are immersed, and returned to their usual habits.

In the IG, the greatest reduction of BMIZ was observed in the first 6 months but did not continue to decrease at 12 months. However, in the subgroup of obese children the BMIZ continued improving after 6 months. This effect has been reported in other study that used a website to send information to parents [35]. The question remains as to how much longer the effect of the intervention can last, especially in obese children who continue to improve their BMIZ.

Concerning obese children of the CG who maintained their BMIZ (rather than increasing it), we propose that parents could have changed to healthier habits after receiving a letter with interpretation of their children's anthropometric measurements and health care recommendations. It has been reported that regular anthropometric measures can improve or reduce BMI [36].

Regardless of the specific intervention, it seems that a crucial element of our strategy was that it targeted parents, who play a central role in promoting healthy habits within the family [33]. Similar results have been reported by Haerens et al [33] in a 2-year follow-up study using electronic means (but no website information) who found that parental involvement was a key factor to maintain BMIZ values. In other studies, in which information was sent to parents via a website, a direct association between the number of visits to the website and reduction in BMIZ was observed [15,16,37]. In this context, parents who most often consult the website are presumably the most informed and implement more changes in their children's eating habits and physical activity.

In our study, the parents of 91.2% (206/226) of IG children received text messages, the highest coverage of all the activities of the strategy. Each of the phone messages summarizes a topic that was extensively explained on the website. The use of mobile phones could play an important role in making successful interventions because other studies have found stronger communication effects using text messages versus the Internet [38]. Additionally, the use of mobile phones has achieved great efficacy at promoting improvements in maternal and child health in resource-poor countries such as India [39], and perhaps this device could be useful in crowded cities with scarce means of transportation.

The main limitations arise in our study: (1) The assignment of participating schools was not randomized. Despite this problem, the BMIZ of the schoolchildren of the IG and CG were comparable at baseline and the effect of the intervention was assessed with the BMIZ change (Δ) during the course of the study. (2) The study participation was approximately 50% and the reasons for nonparticipation should be explored in future studies. Nevertheless, parents of the 2 groups who agreed to participate remained in the study in relatively high proportion at 12 months. (3) The changes in dietary intake and physical activity of participants are not included in this report, and we can only assume that the changes in BMIZ of children were due to the improvement of these habits. (4) The effect of each of the components on BMIZ modification was not evaluated separately in this study because neither in-person activities nor the website consultation showed 100% parental participation. However, the positive effect observed in both normal weight children and obese children of IG possibly is due to the intervention.

The following are the strengths of this study: (1) The study sent information to parents by several channels, in person, through children, via the Internet, and via mobile phones. The few in-person sessions allowed parents a more efficient use of their time in a large, complex city in which in-person meetings are difficult. (2) The follow-up lasted 12 months, longer than in other studies in which interventions last only a few months. (3) Although the intervention had multiple components, it was affordable. (4) The intervention was not only addressed to the obese population but also involved schoolchildren of all BMI classes. This allowed us to determine that, although the effect of the intervention was greater in children with obesity, children with normal weight maintained their BMI, indicating that exposure to obesity risk factors could be decreased with this type of intervention.

Fighting overweight and obesity is a complex undertaking, and some intervention studies in schools have shown effectiveness in restraining but not reducing obesity [34]. Other studies have achieved behavioral changes without reducing BMIZ, and further studies focused on prevention rather than obesity treatment are currently required.

Although the results of this study cannot be generalized, the intervention is promising for implementation of strategies at a distance. Sending information via the Web and smarter mobile phones could encourage better eating and physical activity habits with the aim of preserving or improving children's BMI. In large cities where mobility is difficult but the use of the aforementioned devices has increased [11], it could be useful to take advantage of this kind of intervention for the well-being of schoolchildren.

Lastly, serious and accessible information to promote healthy habits at home, which comes from the school itself, supported by a health institution such as HIMFG, always will be a counterbalance to the commercial information given by the media.

The comprehensive intervention combining in-person activities with Web-based information and mobile phone messaging maintained the BMIZ of normal weight children and decreased the BMIZ of children with obesity. Thus, this can be an affordable alternative to promote health changes in children at the household level.

Acknowledgments

Authors' Contributions

JVG participated in the design and coordination of the study as well as supervision of the fieldwork and conducting the statistical analysis. MKK participated in the design and coordination of the study and critically revised the manuscript. XD participated in conducting the statistical analysis. SFH participated in the conception and design of the research question and provided critical comments on the manuscript. All authors were involved in drafting the manuscript. All of them read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Selected screenshots of the materials used in the intervention.

[[PNG File, 618KB - jmir_v18i6e174_app1.png](#)]

Multimedia Appendix 2

Selected screenshot of the website.

[[PNG File, 289KB - jmir_v18i6e174_app2.png](#)]

Multimedia Appendix 3

Selected screenshots of examples of messages sent to parents' mobile phone.

[[PNG File, 157KB - jmir_v18i6e174_app3.png](#)]

References

1. Freedman DS, Khan LK, Dietz WH, Srinivasan SR, Berenson GS. Relationship of childhood obesity to coronary heart disease risk factors in adulthood: the Bogalusa Heart Study. Pediatrics 2001 Sep;108(3):712-718. [Medline: 11533341]
2. Owen CG, Whincup PH, Orfei L, Chou Q, Rudnicka AR, Wathern AK, et al. Is body mass index before middle age related to coronary heart disease risk in later life? Evidence from observational studies. Int J Obes (Lond) 2009 Aug;33(8):866-877 [FREE Full text] [doi: 10.1038/ijo.2009.102] [Medline: 19506565]
3. Messiah SE, Arheart KL, Luke B, Lipshultz SE, Miller TL. Relationship between body mass index and metabolic syndrome risk factors among US 8- to 14-year-olds, 1999 to 2002. J Pediatr 2008 Aug;153(2):215-221. [doi: 10.1016/j.jpeds.2008.03.002] [Medline: 18534237]
4. Gutierrez J, Rivera-Dommarco J, Shamah-Levy T, Villalpando-Hernandez S, Franco A, Cuevas-Nasu L, et al. Encuesta Nacional de Salud y Nutrición 2012. Resultados Nacionales. Cuernavaca, Mexico: Instituto Nacional de Salud Pública; 2012.
5. Lobstein T, Jackson-Leach R, Moodie ML, Hall KD, Gortmaker SL, Swinburn BA, et al. Child and adolescent obesity: part of a bigger picture. Lancet 2015 Jun 20;385(9986):2510-2520 [FREE Full text] [doi: 10.1016/S0140-6736(14)61746-3] [Medline: 25703114]
6. Martinez JA. Body-weight regulation: causes of obesity. Proc Nutr Soc 2000 Aug;59(3):337-345. [Medline: 10997649]
7. Flores-Huerta S, Klünder-Klünder M, Medina-Bravo P. Elementary school facilities as an opportunity to prevent overweight and obesity in childhood. Bol Med Hosp Infant Mex. 2008. (6) p. 626-638 URL: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-11462008000600018 [WebCite Cache ID 6iJM4iCuA]
8. Prentice AM, Jebb SA. Obesity in Britain: gluttony or sloth? BMJ 1995 Aug 12;311(7002):437-439 [FREE Full text] [Medline: 7640595]
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. Nguyen B, Kornman KP, Baur LA. A review of electronic interventions for prevention and treatment of overweight and obesity in young people. Obes Rev 2011 May;12(5):e298-e314. [doi: 10.1111/j.1467-789X.2010.00830.x] [Medline: 21348921]
11. Manzoni GM, Pagnini F, Corti S, Molinari E, Castelnuovo G. Internet-based behavioral interventions for obesity: an updated systematic review. Clin Pract Epidemiol Ment Health 2011 Mar;7:19-28 [FREE Full text] [doi: 10.2174/1745017901107010019] [Medline: 21552423]
12. Waters E, de SA, Hall BJ, Brown T, Campbell KJ, Gao Y, et al. Interventions for preventing obesity in children. Cochrane Database Syst Rev 2011(12):CD001871. [doi: 10.1002/14651858.CD001871.pub3] [Medline: 22161367]
13. Monasta L, Batty GD, Macaluso A, Ronfani L, Lutje V, Bavcar A, et al. Interventions for the prevention of overweight and obesity in preschool children: a systematic review of randomized controlled trials. Obes Rev 2011 May;12(5):e107-e118. [doi: 10.1111/j.1467-789X.2010.00774.x] [Medline: 20576004]
14. Brown T, Summerbell C. Systematic review of school-based interventions that focus on changing dietary intake and physical activity levels to prevent childhood obesity: an update to the obesity guidance produced by the National Institute for Health and Clinical Excellence. Obes Rev 2009 Jan;10(1):110-141. [doi: 10.1111/j.1467-789X.2008.00515.x] [Medline: 18673306]
15. Delamater AM, Pulgaron ER, Rarback S, Hernandez J, Carrillo A, Christiansen S, et al. Web-based family intervention for overweight children: a pilot study. Child Obes 2013 Feb;9(1):57-63 [FREE Full text] [doi: 10.1089/chi.2011.0126] [Medline: 23308372]
16. White MA, Martin PD, Newton RL, Walden HM, York-Crowe EE, Gordon ST, et al. Mediators of weight loss in a family-based intervention presented over the internet. Obes Res 2004 Jul;12(7):1050-1059. [doi: 10.1038/oby.2004.132] [Medline: 15292468]

17. Comisión Federal de Telecomunicaciones (COFETEL). Scribd. 2012. Informe de resultados 2006-2012 URL: https://es.scribd.com/doc/179740145/INFORME-DE-RESULTADOS-2006-2012-CFT [accessed 2016-06-09] [WebCite Cache ID 6i8kzTCUM]

18. Secretaría de Salud. SEP. México: Diario Oficial de la Federación; 2010. ACUERDO mediante el cual se establecen los lineamientos generales para el expendio o distribución de alimentos y bebidas en los establecimientos de consumo escolar de los planteles de educación básica URL: https://www.sep.gob.mx/work/models/sep1/Resource/635/3/images/acuerdo_lin.pdf [accessed 2016-06-09] [WebCite Cache ID 6i8l9tpFB]

19. Flores-Huerta S, Klünder-Klünder M, Medina-Bravo P. General guidelines established for the sale or distribution of food and beverages consumed in basic education establishments: critical analysis of the AGREEMENT. Bol Med Hosp Infant Mex. 2011. (1) p. 63-70 URL: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-11462011000100010&lng=en&nrm=iso&tlng=en [WebCite Cache ID 6iJM7zuZz]

20. Reedy J, Krebs-Smith SM. Dietary sources of energy, solid fats, and added sugars among children and adolescents in the United States. J Am Diet Assoc 2010 Oct;110(10):1477-1484 [FREE Full text] [doi: 10.1016/j.jada.2010.07.010] [Medline: 20869486]

21. Vereecken CA, Todd J, Roberts C, Mulvihill C, Maes L. Television viewing behaviour and associations with food habits in different countries. Public Health Nutr 2006 Apr;9(2):244-250. [Medline: 16571179]

22. Hendrie GA, Coveney J, Cox DN. Defining the complexity of childhood obesity and related behaviours within the family environment using structural equation modelling. Public Health Nutr 2012 Jan;15(1):48-57. [doi: 10.1017/S1368980011001832] [Medline: 21806870]

23. Linabery AM, Nahhas RW, Johnson W, Choh AC, Towne B, Odegaard AO, et al. Stronger influence of maternal than paternal obesity on infant and early childhood body mass index: the Fels Longitudinal Study. Pediatr Obes 2013 Jun;8(3):159-169 [FREE Full text] [doi: 10.1111/j.2047-6310.2012.00100.x] [Medline: 23042783]

24. Melendez G. Factores asociados con sobrepeso y obesidad en el ambiente escolar. Mexico: Medica Panamericana; 2008.

25. Gidding SS, Lichtenstein AH, Faith MS, Karpyn A, Mennella JA, Popkin B, et al. Implementing American Heart Association pediatric and adult nutrition guidelines: a scientific statement from the American Heart Association Nutrition Committee of the Council on Nutrition, Physical Activity and Metabolism, Council on Cardiovascular Disease in the Young, Council on Arteriosclerosis, Thrombosis and Vascular Biology, Council on Cardiovascular Nursing, Council on Epidemiology and Prevention, and Council for High Blood Pressure Research. Circulation 2009 Mar 3;119(8):1161-1175 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.109.191856] [Medline: 19255356]

26. Secretaria de Salud. DOF. Mexico; 2013. Norma Oficial Mexicana NOM-043-SSA2-2012: Servicios básicos de salud. Promoción y educación para la salud en materia alimentaria. Criterios para brindar orientación URL: http://dof.gob.mx/nota_detalle.php?codigo=5285372&fecha=22/01/2013 [accessed 2016-06-09] [WebCite Cache ID 6iJMCaXPv]

27. Rivera JA, Muñoz-Hernández O, Rosas-Peralta M, Aguilar-Salinas CA, Popkin BM, Willett WC, Comité de Expertos para las Recomendaciones. Beverage consumption for a healthy life: recommendations for the Mexican population. Salud Publica Mex 2008;50(2):173-195 [FREE Full text] [Medline: 18372998]

28. World Health Organization. Physical status: the use and interpretation of anthropometry. Geneva: World Health Organization; 1995.

29. World Health Organization. 2007. Growth Reference Data for School-aged Children and Adolescents of 5-19 Years URL: http://www.who.int/growthref/en/ [accessed 2016-06-05] [WebCite Cache ID 6eoDj4w0K]

30. Klünder-Klünder M, Flores-Huerta S. Waist circumference values according to height percentiles: a proposal to evaluate abdominal obesity in Mexican children and adolescents between 6 and 16 years of age. Arch Med Res 2011 Aug;42(6):515-522. [doi: 10.1016/j.arcmed.2011.09.004] [Medline: 21925222]

31. Maggio AB, Saunders GC, Gal-Duding C, Beghetti M, Martin XE, Farpour-Lambert NJ, et al. BMI changes in children and adolescents attending a specialized childhood obesity center: a cohort study. BMC Pediatr 2013;13:216 [FREE Full text] [doi: 10.1186/1471-2431-13-216] [Medline: 24369093]

32. Doyle AC, Goldschmidt A, Huang C, Winzelberg AJ, Taylor CB, Wilfley DE. Reduction of overweight and eating disorder symptoms via the Internet in adolescents: a randomized controlled trial. J Adolesc Health 2008 Aug;43(2):172-179 [FREE Full text] [doi: 10.1016/j.jadohealth.2008.01.011] [Medline: 18639791]

33. Haerens L, Deforche B, Maes L, Stevens V, Cardon G, De Bourdeaudhuij I. Body mass effects of a physical activity and healthy food intervention in middle schools. Obesity (Silver Spring) 2006 May;14(5):847-854. [doi: 10.1038/oby.2006.98] [Medline: 16855194]

34. Kain J, Concha F, Moreno L, Leyton B. School-based obesity prevention intervention in Chilean children: effective in controlling, but not reducing obesity. J Obes 2014;2014:618293 [FREE Full text] [doi: 10.1155/2014/618293] [Medline: 24872892]

35. Williamson DA, Walden HM, White MA, York-Crowe E, Newton RL, Alfonso A, et al. Two-year internet-based randomized controlled trial for weight loss in African-American girls. Obesity (Silver Spring) 2006 Jul;14(7):1231-1243. [doi: 10.1038/oby.2006.140] [Medline: 16899804]

36. Thompson JW, Card-Higginson P. Arkansas' experience: statewide surveillance and parental information on the child obesity epidemic. Pediatrics 2009 Sep;124 Suppl 1:S73-S82. [doi: 10.1542/peds.2008-3586J] [Medline: 19720670]

37. An J, Hayman LL, Park Y, Dusaj TK, Ayres CG. Web-based weight management programs for children and adolescents: a systematic review of randomized controlled trial studies. ANS Adv Nurs Sci 2009;32(3):222-240. [doi: 10.1097/ANS.0b013e3181b0d6ef] [Medline: 19707091]

38. van Wier MF, Ariëns GA, Dekkers JC, Hendriksen IJ, Smid T, van Mechelen W. Phone and e-mail counselling are effective for weight management in an overweight working population: a randomized controlled trial. BMC Public Health 2009;9:6 [FREE Full text] [doi: 10.1186/1471-2458-9-6] [Medline: 19134171]

39. IAP. 2015 Aug. World Alliance for Breastfeeding Action: Global IAP Healthphone Initiative, WABA URL: http://www.iap.healthphone.org/#sthash.B346K6RD.dpuf [accessed 2016-06-09] [WebCite Cache ID 6i8nkpWm8]

Abbreviations

BMI: body mass index
BMIZ: body mass index *z*-score
CG: control group
IG: intervention group
HIMFG: Federico Gomez Children's Hospital of Mexico
UNAM: National Autonomous University of Mexico

XSL·FO
RenderX

Original Paper

How a Fully Automated eHealth Program Simulates Three Therapeutic Processes: A Case Study

Marianne T. S Holter[1], Mphil; Ayna Johansen[1,2], PhD; Håvar Brendryen[1], PhD

[1]The Norwegian Centre for Addiction Research, Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway
[2]Centre for the Study of Mind in Nature, Faculty of Humanities, University of Oslo, Oslo, Norway

Corresponding Author:
Marianne T. S Holter, Mphil
The Norwegian Centre for Addiction Research
Institute of Clinical Medicine, Faculty of Medicine
University of Oslo
Postboks 1039 Blindern
Oslo, 0315
Norway
Phone: 47 93 62 30 61
Fax: 47 23 36 89 86
Email: m.t.s.holter@medisin.uio.no

Abstract

Background: eHealth programs may be better understood by breaking down the components of one particular program and discussing its potential for interactivity and tailoring in regard to concepts from face-to-face counseling. In the search for the efficacious elements within eHealth programs, it is important to understand how a program using lapse management may simultaneously support working alliance, internalization of motivation, and behavior maintenance. These processes have been applied to fully automated eHealth programs individually. However, given their significance in face-to-face counseling, it may be important to simulate the processes simultaneously in interactive, tailored programs.

Objective: We propose a theoretical model for how fully automated behavior change eHealth programs may be more effective by simulating a therapist's support of a working alliance, internalization of motivation, and managing lapses.

Methods: We show how the model is derived from theory and its application to *Endre*, a fully automated smoking cessation program that engages the user in several "counseling sessions" about quitting. A descriptive case study based on tools from the intervention mapping protocol shows how each therapeutic process is simulated.

Results: The program supports the user's working alliance through alliance factors, the nonembodied relational agent *Endre* and computerized motivational interviewing. Computerized motivational interviewing also supports internalized motivation to quit, whereas a lapse management component responds to lapses. The description operationalizes working alliance, internalization of motivation, and managing lapses, in terms of eHealth support of smoking cessation.

Conclusions: A program may simulate working alliance, internalization of motivation, and lapse management through interactivity and individual tailoring, potentially making fully automated eHealth behavior change programs more effective.

Introduction

"Black boxes," or poorly described programs, have long been a criticism of the eHealth field [1-4], and effective program components across individual interventions are still largely unknown [5]. To address this problem, assumed mechanisms should be adequately described and put in a theoretical context [6]. This would build well-founded hypotheses for active program ingredients. Theoretically founded hypotheses may be especially useful in fully automated programs because automation standardize the therapy that is given. The standardization allows for program elements to be described in

detail and investigated empirically, free from human variations and with a large degree of reliability. Investigating eHealth programs in light of counseling theories may increase our understanding of how such programs work [6]. In this paper, we will break down the components of an eHealth program and discuss its potential for interactivity and tailoring in terms of common concepts from face-to-face counseling. We hypothesize that simulating the therapeutic processes of supporting a working alliance [7,8], internalized motivation [9], and lapse management [10] simultaneously may be important to optimize behavior change.

According to Riley and colleagues, traditional health behavior change theories are static and linear in nature, and therefore, do not take advantage of the potential involved with interactive eHealth interventions [6]. eHealth interventions are not necessarily static or linear, as they can follow individual users and respond with tailored output to their immediate and previous responses. This enables dynamic adjustment of the intervention delivered, and theories from face-to-face counseling may therefore be more suited to understand eHealth interventions' effective ingredients [6]. In this paper, therefore, we examine *Endre*, a fully automated program for smoking cessation that uses a fictional "therapist" to conduct tailored "counseling" sessions with the user.

Within eHealth-assisted behavior change, there is a growing interest in the concept of a working alliance [11-22], which is found essential in face-to-face counseling [7,8]. The alliance is commonly defined as an emotional bond, as well as agreement on task and goal [7]. It can also be described as therapist processes—such as empathy, warmth, and genuineness, establishing a collaborative framework and offering support and guidance [23]. A strong alliance facilitates client processes that are central to therapy-assisted behavior change, such as expectancies, intentions, motivation, hope, openness, trust, commitment, satisfaction, and a changing view of the self [23]. It may be possible to develop a working alliance to a fully automated program [12,21,22], but so far, there are only a few examples of programs designed to support a working alliance [11,12]. Likewise, motivational interviewing (MI) [24] is considered an effective method to motivate client change in counseling [25]. The effectiveness of MI has been linked to its ability to influence 3 basic psychological needs, including competence, relatedness, and autonomy [26,27]. By supporting

these needs, external motivation, a weak form of motivation characterized by performing an activity to gain an external reward or avoid an external punishment, can become internalized. This means the activity is performed because the individual accepts it as an important step toward a personally valued goal [26], improving self-regulation, performance, and persistence [9,27,28]. Although MI is often mentioned as one of several methods in eHealth programs [29-33], only 2 report MI as a main method applied extensively [32,33]. Finally, behavior change is difficult, and even when an individual is motivated and the change is going well, he or she still needs to avoid lapses or setbacks in behavior. If a lapse should occur, the individual needs to react constructively to avoid a complete relapse. Teaching people how to prevent a lapse from becoming a relapse (lapse preparation), and helping them manage lapses (lapse management), is thus important when implementing behavior change [10]. Lapse preparation and lapse management have previously been applied to fully automated eHealth programs [31,34-37], but its effect has not been documented. Each therapeutic process has a unique contribution to the user's change process. Supporting internalization of motivation gives the user strength and persistence in upholding the change [9,27,28]. Helping the user manage lapses keeps him or her from resuming the old behavior after a setback. Finally, supporting a working alliance makes a positive therapy outcome more likely [7,8] (Figure 1).

No published description exists, as far as we know, of a program supporting all 3 processes simultaneously, as proposed in the theoretical model in Figure 1. The aim of this paper is therefore to illustrate this model through a case study of *Endre*, a fully automated smoking cessation program, and to forward a hypothesis of these 3 therapeutic processes as important eHealth elements. We use a focused, descriptive analysis to conceptualize the translation from theory to intervention. The analysis is based on a modified intervention mapping protocol [38], which is a framework for designing and planning health promotion interventions through a taxonomy of mapping tools that can be used to code intervention contents. We use the steps that target process theory, methods, and design integration (steps 2-4) to focus on the 3 therapeutic processes that constitute the working hypothesis of *Endre*. This paper therefore also exemplifies the use of intervention mapping as an approach ideally suited to investigate potentially important elements in the "black box" of eHealth programs.

Figure 1. Different therapeutic processes affect behavior change differently.
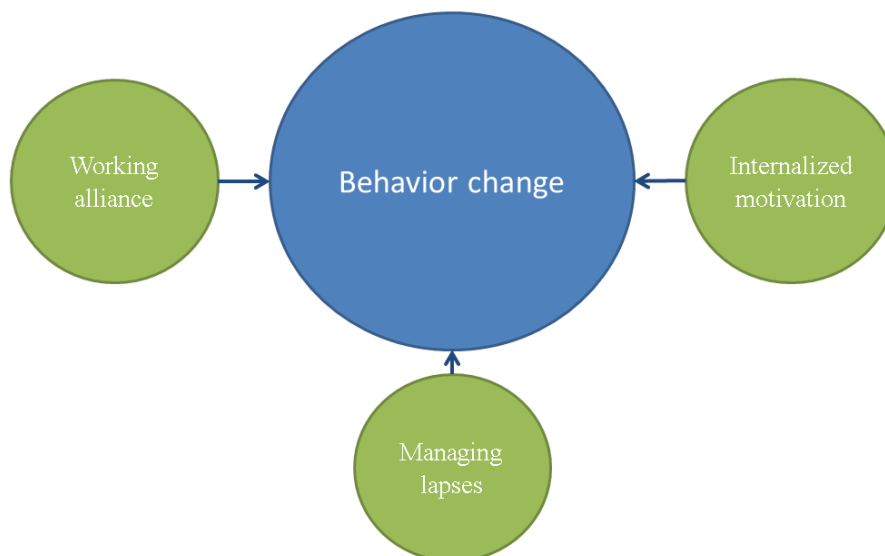
Methods

The Case: Endre

Endre is a fully automated eHealth program for smoking cessation that has evolved from the third author's experience with the smoking cessation program *Happy Ending* [34]. *Endre* has some of the same basic structures as *Happy Ending*. It uses tunneling [18,39], has both pull (Web page) and push elements (e-mails and short message service [SMS] messages), and delivers program materials through the "voice" of a nonembodied relational agent [11]. Importantly, lapse management (with Marlatt's cognitive behavioral model of relapse prevention [10] as methodological counterpart) is a central component of both *Endre* and *Happy Ending*. However, as opposed to *Happy Ending*, which in addition to lapse management consisted of a large number of theoretical and methodological underpinnings [34], the content of *Endre* is centered on 2 other theoretical concepts: internalized motivation (with MI as the methodological counterpart [26,27]) and working alliance (with *alliance factors* [13] as the methodological counterpart).

Endre consists of 26 tunneled [18,39] Web sessions. On registration, users provide their mobile phone number and e-mail address, which prompts receipt of an automatically generated e-mail with a username and password. After the program starts, the user goes through 10 days of preparing to quit with one new session each day, followed by their quit day, which is scheduled on the 11th day. The user must confirm a quit attempt before the program moves on to the follow-up phase. In the follow-up phase, the user gets one new session the first 3 days, then 2 new sessions every week for the first 4 weeks, and finally one new session a week for the last 4 weeks. The program ends 8 weeks after the cessation day. Automatically generated e-mails give the user access to each new session through a link. The links are time based, they lead to today's session for that individual user, and one cannot access earlier sessions by clicking on old links. If a user rarely logs on, he or she will only receive the most important missed sessions. An overview of the themes for each session can be viewed in Multimedia Appendix 1.

Endre provides no additional human support. Most sessions involve user interactivity, requesting input from the user (see screenshots below for examples). We anticipate that an adult, typical user with average reading abilities may spend 4-6 minutes on each session. The user receives synchronous and immediate feedback on input. The lapse management component of *Endre* is based on the lapse management component in *Happy Ending* [34] and consists of daily SMS messages that are sent out to users who have quit, asking them if they have been smoke free that day. If the user reports a lapse, he or she gets access to a special, Web-based session intended to help the user recover from the lapse (Multimedia Appendix 2). This special session can be accessed whenever and for as many times as necessary.

Analytic Procedure

We describe how a counselor's support of a working alliance, internal motivation, and lapse preparation and management are simulated in *Endre* by using selected steps from the intervention mapping protocol (steps 2, 3, and 4) [38]. Intervention mapping is well suited for describing process simulation because it can be applied to understand the program construction. Furthermore, the necessary information for an intervention mapping analysis was readily available, as *Endre* was developed using intervention mapping. Intervention mapping is conventionally used to describe everything in a program [29,30,40-50]. Contrarily, we use it in a focused way to describe only the elements that are relevant to our hypothesis of important program elements. The intervention mapping tools are thus used for an analysis consisting of 2 parts: First going from general therapeutic process to theoretical operationalization suiting the context of this program; and second, going from theoretical operationalization to simulation in specific program elements.

First, we describe how supporting a working alliance, internalized motivation, and lapse preparation and management are operationalized in *Endre*'s theoretical change model (step 2 in intervention mapping [38]). In the change model, the changes necessary to quit smoking by means of *Endre* are

described and displayed in a matrix. In the intersecting cells of the matrix, the operationalization of each therapeutic process is described in a list of change objectives. That is, each change objective shows how one aspect of one of the therapeutic processes is operationalized for the purpose of the intervention (that the user quits smoking and stays smoke free with *Endre*) and its context (a fully automated program). An analytic text accompanies the change model to describe how the 3 processes are represented in the change model. The change model that was used for the development of *Endre* (Multimedia Appendix 3) is simplified to highlight the 3 therapeutic processes, and we use sequential numbering of the change objectives instead of conventional intervention mapping-labeling [38] to improve readability outside of the intervention mapping community. The change model operationalizes the abstract and general therapeutic processes. It is therefore the first part of the analysis toward the processes' simulation.

After showing how supporting a working alliance, internalized motivation, and lapse preparation and lapse management are operationalized through change objectives, we describe how the 3 therapeutic processes are simulated through specific program elements (steps 3 and 4 in intervention mapping [38]). The program elements result from combining change objectives with theoretical methods for inducing change (eg, MI, modeling). This second part of the analysis takes the (theoretical) operationalizations of the 3 therapeutic processes and makes them into (practical) simulations through specific program elements.

Results

Operationalization of the Therapeutic Processes in Endre

The operationalization of the therapeutic processes can be viewed in the change model matrix (Table 1). In the matrix, sub-behaviors in quitting, or performance objectives, are crossed with theoretical constructs, or personal determinants , believed to be causing or influencing the behavior. Each therapeutic process is represented within the model either as a personal determinant or a performance objective. The personal determinants and performance objectives intersect in cells containing change objectives , which specify how each personal determinant must change for the individual to be equipped to do each performance objective.

Working alliance and internalized motivation are operationalized as personal determinants, whereas behavior maintenance through lapse preparation and lapse management is operationalized as a performance objective. Having a working alliance to the program is not a necessary psychological process for quitting smoking in general. It might, however, be an important process for quitting smoking with the help of *Endre*, if one assumes that a successful simulation of supporting a working alliance can have the same benefits for therapy outcome in a fully automated program as it has in face-to-face counseling [7,8]. Though a working alliance can be an important psychological process for quitting smoking with *Endre*, internalized motivation is an important psychological process for succeeding in quitting smoking at all. In the model, internalized motivation is separated into the underlying personal determinants relatedness, competence, and autonomy; the 3 "needs" that influence the internalization of motivation [9]. Competence is itself separated into 2 personal determinants: skills and self-efficacy. As with competence, relatedness is also separated into 2 personal determinants: relatedness to social network and working alliance. Working alliance, or relatedness to the program, is included under relatedness because a positive counseling relationship can also support the client's (or user's) need for relatedness [27]. In contrast, behavior maintenance through lapse prevention and lapse management is operationalized in the change model as a performance objective, meaning that managing lapses in a constructive way is considered an important subgoal for succeeding in quitting smoking. The change objectives belonging to each therapeutic process is the operationalization of that process for the purpose of this program.

Table 1. Modified change model.[a]

Performance objectives	Personal determinants				
	Internalized motivation (therapeutic process 2)				
	Relatedness		Competence		Autonomy
	Working alliance (therapeutic process 1)	To social network	Skills	Self-efficacy	
1. Decide to quit smoking and plan how to do it.	1. Experience the program as a social actor [11]. 2. Experience the program as accessible, helpful, empathic, and trustworthy [13]. 3. Be aware of one's influence on program content [13]. 4. Understand how to use the program and do the exercises [13].	5. Make a public commitment to the quit attempt. 6. Choose a "support person" from one's personal network.	7. (1) Identify personal smoking cues and (2) be able to detect smoking urges early. 8. Make an action and coping plan for the quit attempt. 9. (1) Identify one's high-risk situations, and (2) make an action and coping plan for handling them [24].	10. Believe it to be possible to quit smoking and stay smoke free [13]. 11. Be confident in one's ability to execute the action and coping plan [24].	12. Commit personally to the quit attempt and know one's personal reasons for doing so [24]. 13. Decide whether or not to (1) make a public commitment and (2) engage a "support person." 14. Choose how to make the quitting plan (by oneself or a more guided version) [24]. 15. Combine the advice of the program with one's own style and preferences [13].
2. Initiate the quit attempt and stay smoke free for the first 3 days.	16. Experience the program as: (1) responsive, sensitive, and adjustable for emerging needs and (2) suiting one's own preferences and style [13].	17. Ask the "support person" for practical assistance and emotional support as needed [51].	18. Implement action and coping plan. 19. Get rid of remaining cigarettes and smoking accessories. 20. Withstand cravings and cope with withdrawal symptoms.	21. Be confident in one's ability to stay smoke free the first 3 days [24].	22. Revise the action and coping plan if needed. 23. Decide whether or not to get rid of remaining cigarettes, or whether to make the cigarettes less accessible. 24. Decide to what degree, when and how the "support person" is needed.
3. Establish a smoke-free lifestyle (from day 4 and onward).	25. Continue with the program for as long as needed, even after a period of program disengagement ("rupture prevention and repair") [13].	Same as the above (change objective 17).	26. Identify and counteract thought patterns that could lead to a (re)lapse [52]. 27. Follow plans for high-risk situations.	28. Be confident in (1) one's ability to continue being smoke free, and (2) one's ability to stay smoke free in the long run.	29. Attribute success in the cessation attempt internally.
4. Maintain the behavior by managing lapses constructively (therapeutic process 3).	Same as the above (change objectives 16 and 25).	30. Explain the difference between a lapse and a relapse to significant others to gain their continued support.	31. Know the difference between a lapse and a relapse, and how to recover from a lapse. 32. Get rid of any spare cigarettes after a lapse. 33. Resist new urges to smoke.	34. Be confident in one's ability to continue with the quit attempt after a lapse.	35. Know that whether or not to remain smoke free is a matter of one's own choice.

[a]Every cell specifies the theoretical operationalization of one (or several) therapeutic process(es).

XSL•FO
RenderX

Simulation of the Therapeutic Processes in Endre

In this section, we describe how the therapeutic processes are simulated in *Endre*. For each therapeutic process, we present program elements that are involved in the simulation and describe the methods that are used. To support working alliance we adapted MI [24] to a computerized "counselor" who delivers all program material through what we refer to as computerized motivational interviewing (cMI). The "counselor" is called *Endre*, which has a double meaning in Norwegian, being a man's name, as well as literally meaning "to change." Internalized motivation is primarily supported through cMI, whereas behavior maintenance is strengthened with a psycho-educative session before the quitting day, as well as a special Web-based session that is made accessible if the user reports a lapse. If the user experiences several lapses, this is recognized by *Endre*, and the content of the session is adjusted accordingly.

Simulation of Working Alliance Support

Working alliance is supported in program elements using a nonembodied relational agent [12], cMI (Multimedia Appendix 4), and dynamic tailoring [42] to convey alliance factors [13]. For the users to experience the program as a social actor [11] (change objective 1) that is accessible, helpful, empathic, and trustworthy [13] (change objective 2), the relational agent [12] *Endre* is used throughout the program. *Endre* is a nonembodied, text-based relational agent that simulates a "counselor" the user "communicates" with. Some key attributes of *Endre* can be found in Textbox 1, and examples of how "he" is represented in the program can be found in Figures 2-6.

Textbox 1. Attributes of the relational agent Endre.

Uses first person tense.

Introduces a new topic for each session.

Asks questions and reflects answers empathically [24].

Uses appropriate greetings and farewells according to time of day [12].

Uses humor [12].

"Remembers" earlier conversations by explicitly referring to them or implicitly adjusting program content.

To further support a working alliance, users are allowed to influence the program content [13] (change objective 3). This is a way of "negotiating" goals [13] and is done in the first session (Figure 2). After *Endre* has presented the program plan, the user is asked to choose a topic he or she considers important when quitting. On the subsequent page, *Endre* assures the user that "he" will make time for this topic during the course of the program. The user's topic is visited 2 times during the program.

To build a working alliance to the user, it is also necessary for him or her to receive guidance in how to use the program [13] (change objective 4). *Endre* provides guidance to the user, for example by explaining how new sessions are made available and how the user can log onto them. In addition, new program exercises are demonstrated by four fictional "quitters" (Figure 3).

Working alliance is further strengthened if the program is experienced as responsive, sensitive, adjustable for emerging needs, and suiting one's own preferences and style [13] (change objective 16). To address this, *Endre* has a *flexible session manager* (Textbox 2) that adjusts the total number of sessions to user behavior. This means that a user who does not log on to the program every time a new session is available will only receive the most important sessions, limiting the total number of sessions for that particular user.

Textbox 2. Flexible session manager.

Ensures that a user who has missed several sessions receives the most important session of the ones he or she has missed.

A user that seldom logs on will only get the most important sessions of the program.

We developed a set of rules that decides what session the user will get next (ie, the most important sessions), based on:

the program plan,

which sessions the user has already logged on to,

rules that categorize the sessions as either high priority (all users must go through these) or low priority (the user only receives these if he or she has done the high-priority sessions thus far).

A user who has missed several sessions first receives those that are categorized as high priority.

If the user has logged on to all high-priority sessions, he or she receives low-priority sessions that address (in the following order): skills, self-efficacy, relatedness, and autonomy.

A final program aspect supporting working alliance is a "mini motivation intervention" (Textbox 3), consisting of SMS messages and intended to prevent program disengagement (alliance "rupture" [13]) (change objective 25).

Textbox 3. Mini motivation intervention.

Before quitting day:

If the user misses one session, nothing happens.

If the user misses 2 sessions, he or she gets an SMS message from *Endre*, reminding him or her to log on.

If the user misses 3 sessions, he or she gets an SMS message where *Endre* normalizes having second thoughts and recommends logging on to the program.

If the user misses a fourth session, nothing happens.

If the user misses a fifth session, he or she gets a final SMS message where *Endre* appeals to the "healthy part" of the user to log on.

After quitting day:

After quitting day, there is no intervention if the user does not log on to the Web page.

A part of the lapse management system is that the user every evening receives an SMS message, asking if he or she has been smoke free. If the user does not answer the SMS message, he or she will receive up to 3 extra SMS messages encouraging him or her to answer.

Figure 2. Choosing a topic ("negotiating" goals).

Figure 3. The 4 "quitters" demonstrate how to do the program exercises and model how to combine Endre's advice with one's own personal style.

Figure 4. Eliciting self-efficacy change talk through a confidence ruler.

Figure 5. Endre has asked the user to choose a "support person" for her quit attempt, and the user has answered that he or she wants to quit without any help.

Figure 6. Handling sustain talk and discord.
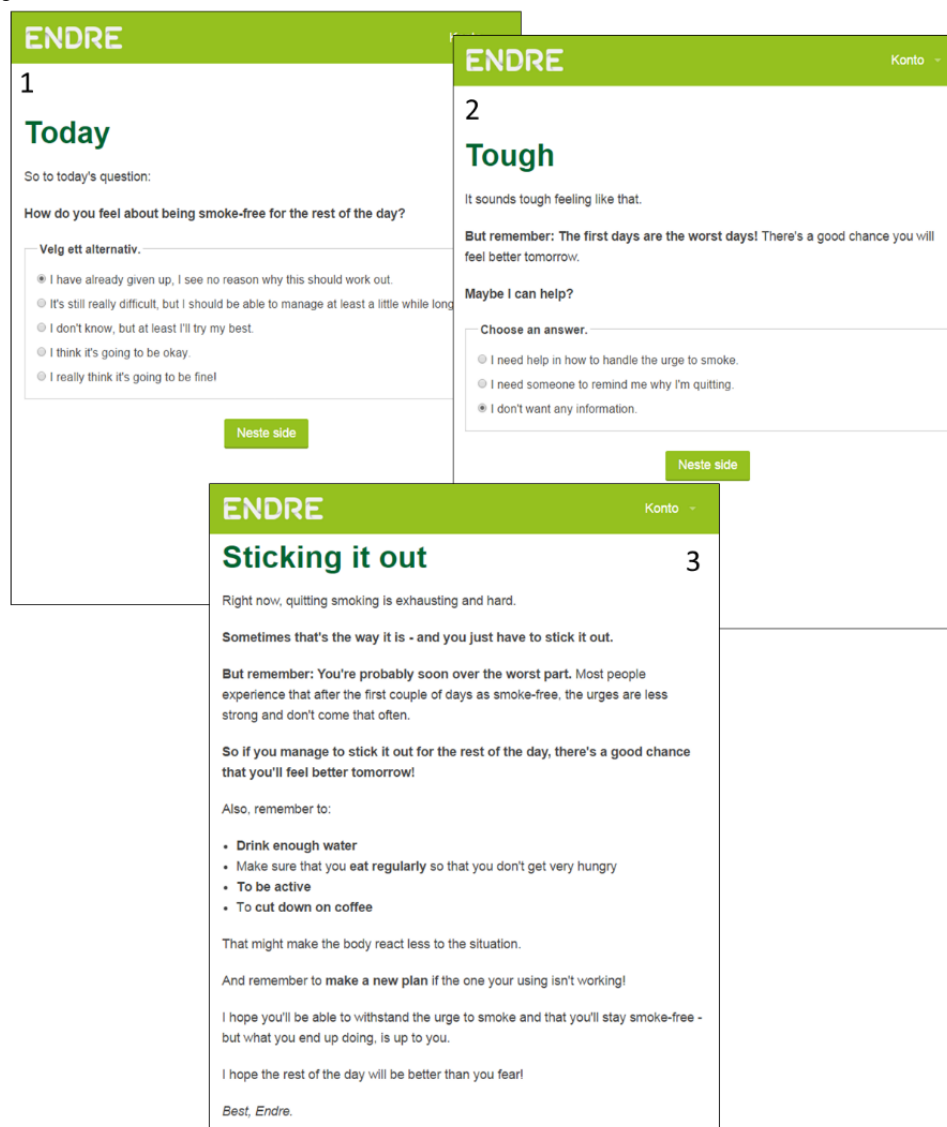
Simulation of Internalized Motivation Support

Internalized motivation is achieved through *Endre* strengthening the user's autonomy, competence, and relatedness [9]. Relatedness is partly supported through building a working alliance between the user and the program, as described in the previous section. The other part of relatedness, relatedness to social network, is strengthened through helping the user find support in the people surrounding him or her. This is done by advising the user to recruit a "support person" from his or her social network (change objective 6), advising him or her to make the quit attempt public (change objective 5), and guiding the user in how to make their "support person" have the greatest positive impact on his or her quit attempt (change objective 17). Figure 5 is from the session where *Endre* advices the user to choose a "support person," showing what happens when the user does not want to follow *Endre* 's advice. Another way in which *Endre* supports the user's relatedness to his or her social network is effectuated if the user reports a lapse after he or she has quit. *Endre* then asks if this lapse may affect the user's relationship to his or her social network. If the user answers yes, *Endre* offers help to ensure the social network's continued support for the quit attempt (change objective 30). All advice is given using cMI (Multimedia Appendix 4) and dynamic tailoring [42].

Autonomy is supported in program elements using cMI, dynamic tailoring [42], and modeling [38]. One way *Endre* supports autonomy is by asking for permission before giving any information or advice (Textbox 4). This is a way of acknowledging that the user chooses what information to receive. Asking for permission is relevant to change objectives 13, 15, 23, and 24.

Textbox 4. Asking for permission.

Endre requires one of two actions from the user before giving any information or advice:

Hide/show text: The user can choose whether or not to click on the question (eg, "do you want me to tell you about…") to reveal the information. An example can be viewed in Figure 3.

Question + multiple choice yes/no: The user must answer yes or no when *Endre* asks for permission to give information or advice; if the user answers yes, the information is revealed on the next page. For example, one session starts by *Endre* introducing today's topic, and then asking the user whether he or she thinks this sounds okay. If the user answers yes, the session continues. If the user answers no, the session is ended.

A second way in which autonomy is supported is through handling *sustain talk* (reasons for smoking) and *discord* (dissatisfaction with therapy) [24] respectively. Sustain talk and discord may be expressed by the user at select places in the program through multiple-choice alternatives. The fact that expressing sustain talk or discord is allowed (even when it goes against the program) communicates respect for the user's autonomy. If sustain talk or discord is expressed, *Endre* repeats the user's feelings empathically, and then, depending on the situation, asks more questions, normalizes, offers help, or changes the topic [24]. Handling sustain talk and discord is relevant to change objectives 13, 15, 23, 24, and 35. An example of how sustain talk or discord may be expressed and how it is handled can be seen in Figure 6. This is from the user's second day as smoke free. On page 1, *Endre* asks the user how he or she feels about staying smoke free for the rest of the day. The example shows the user choosing the statement representing the lowest degree of self-efficacy; so low that it qualifies as sustain talk. On page 2, *Endre* offers help. The user chooses that he or she does not want any help; this can be seen as dissatisfaction with the program, or discord. On page 3, *Endre* reflects empathically and normalizes the user's feelings.

A third way *Endre* supports the user's autonomy is through eliciting and reflecting *change talk*, that is, talk arguing toward change [24] (change objective 12). Change talk is the user's autonomous reasons and capacities for quitting and is requested throughout the program. *Endre* repeats the user's change talk and sometimes elaborates on it. For example, in one session, *Endre* asks the user for his or her most important reason for wanting to become smoke free (eliciting change talk). *Endre* repeats the user's most important reason on the next page (reflecting change talk). Asking for permission, handling sustain talk and discord, and eliciting change talk is achieved through cMI and dynamic tailoring [42], and details of these applications can be viewed in Multimedia Appendix 4.

A fourth and final way in which *Endre* supports autonomy is through modeling [38]. In the program, 4 fictional "quitters" model autonomy by illustrating how to combine the advice of the program with one's own style and preferences (change objective 15). The 4 "quitters" are of different gender, age, socioeconomic status, cultural background, and *smoking profiles* [53]. The "quitters" answer *Endre*'s questions and tasks in ways that suit their situation and personality. An example of this application can be viewed in Figure 3. This screenshot is from the session for making a cessation plan, where *Endre* asks the user what he or she needs to do the day before quitting. By clicking on the names of the 4 fictional "quitters," the user may read "their" answers.

Autonomy is supported through asking for permission before giving advice, handling sustain talk and discord respectfully, eliciting and reflecting change talk, and modeling how to combine the program's advice with one's own preferences and style. Competence is supported through increasing the user's quit-related skills and increasing his or her self-efficacy for quitting. Skills can be acquired through the general information and advice that *Endre* gives, as well as through program exercises. For example, before quitting day, *Endre* asks the user to spend a few days thinking about what precedes his or her smoking—what are his or her *smoking cues*. After a few days, *Endre* asks the user for these smoking cues. This teaches the user to be attentive to what triggers the urge to smoke. The advices and exercises that *Endre* gives are based on self-monitoring of behavior, counter-conditioning, active learning, goal setting, planning coping responses, and implementation intentions [38], always communicated using cMI (Multimedia Appendix 4).

Whereas skills are supported through information, advice, and exercises, self-efficacy is supported through cMI techniques, in combination with dynamic tailoring [42]. The user's self-efficacy is strengthened through "confidence rulers" [24,32]. An example of this application can be found in Figure 4. These screenshots are from the same session as the ones in Figure 6, but showing what happens when the user answers differently. In this example, the user chooses the statement reflecting a quite high degree of self-efficacy. On page 2, *Endre* asks the user to justify why he or she chose that statement over a statement representing a lower degree of self-efficacy. The user types in his or her answer, and on page 3, this statement is reflected back to him or her. The user has argued for change and had the argument reflected back, amplifying the effect [24].

Self-efficacy is also strengthened through affirmations [24], that is, compliments on the user's strengths and accomplishments. For example, in one session, the user is asked if he or she has tried quitting before. If the user answers yes, *Endre* replies that this is a good thing, because the user then has experience that he or she can use to increase the chances of succeeding this time. Turning previous quitting experience into something positive is a way of providing affirmation, supporting self-efficacy, competence, and internal motivation.

Simulation of Lapse Preparation and Lapse Management Support

Behavior maintenance is supported through a psychoeducative session before the user's quit day and a lapse management component after he or she has quit. First, a psychoeducative session on lapses and relapses prepares the user to respond constructively in case of a lapse (change objectives 31 and 35).

XSL•FO

RenderX

In this session, a car puncturing a tire is used as a visual analogy [38] for lapsing and relapsing. The cars can be seen in Figure 7. Car no.1 illustrates the lapse (puncturing the tire), car no.2 illustrates a relapse (giving up and succumbing to negative emotions), car no.3 shows the process of choosing, car no.4 is acting to resume the quit attempt, and car no. 5 illustrates being smoke free again.

In the preparatory session, the user is also presented with an advance organizer [38] of the process of becoming smoke free again after a lapse. The advance organizer has the shape of a circle (Figure 8) displaying the self-regulation loops [54] that can help the user back to being smoke free. First, realize that you are smoking ("innse"), then choose: Keep smoking or keep quitting ("velge"), then act to become smoke free again ("handle"), and finally continue with being smoke free ("fortsett"). The information is given with cMI (Multimedia Appendix 4).

Following up on the preparatory session on lapses and relapses is a lapse management component which is effectuated after the user has confirmed a quit attempt. Every day, the user receives an SMS message asking if she is still smoke free. If the user answers yes, another SMS message compliments the user's accomplishment. If however the user answers no, he or she receives an SMS message with a link to a Web-based lapse management session. The user may access the session through the SMS message; if he or she does not log on via the SMS message, he or she receives the lapse management session when logging on to the program next time. The lapse management

session helps the user make a choice, become smoke free again and learn from the lapse. When logging on to the Web-page, the user is first reintroduced to the car (Figure 7) and the circle (Figure 8). *Endre* then asks if the user has already decided what to do: keep quitting or keep smoking (Figure 9). If the user chooses to keep quitting, *Endre* guides him or her back to being smoke free, helps making a new plan on how to face a similar situation in the future without lapsing, and supports the user's belief in his or her ability to stay smoke free. Figure 9 shows a screenshot from the lapse management session. In this example, *Endre* has asked the user if he or she knows what to do now, and the user has answered that he or she is unsure. On the next page, shown in the screenshot, the user may choose which topic he or she wants *Endre* to start with (the picture does not show the entire page). Asking the user what topic to start with is a way of asking for permission [24], strengthening his or her autonomy and supporting internal motivation. In addition, letting the user influence the program structure influences the working alliance positively [13]. This screenshot shows the main topics that are covered in the lapse management session: reattribution [10], ambivalence [24], the abstinence violation effect [10], and making a choice. Only users who express ambivalence or an abstinence violation effect when asked go through these topics. Multimedia Appendix 2 contains more information on the lapse management component, including a flow chart that shows the different ways in which this session may be built up. Some of the methods that are used are cMI (Multimedia Appendix 4), dynamic tailoring [42], reattribution [55], and cognitive restructuring [56].

Figure 7. Visual analogy for lapsing and resuming the quit attempt.

Car No.1   Car No.2

Car No. 3   Car No. 4   Car No.5
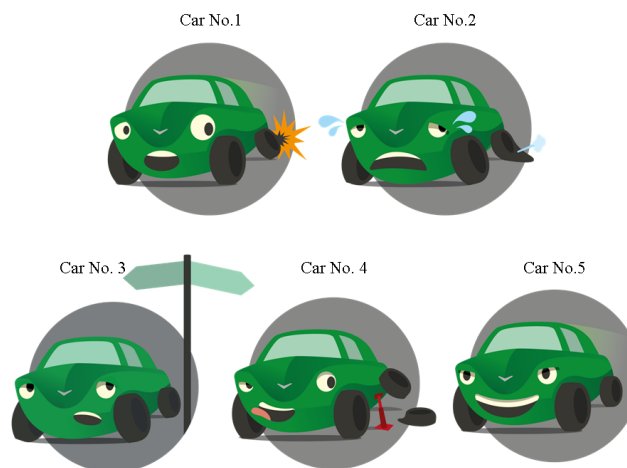
Figure 8. Advance organizer of returning to the quit attempt after a lapse (from top left section): realize ("innse"), choose ("velge"), act ("handle"), and continue ("fortsett").
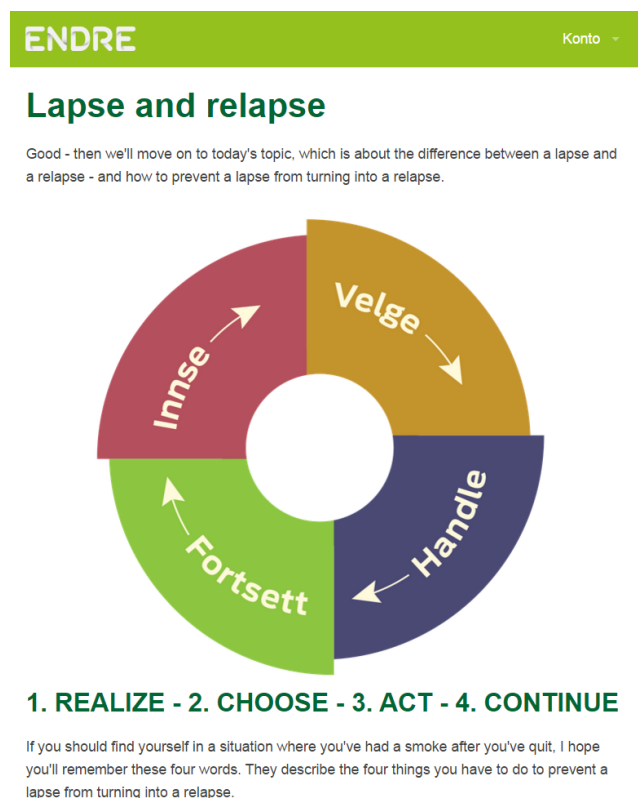
Figure 9. From the lapse management session: the user is unsure of what to do and is asked what topic to begin with.

Discussion

Summary Analysis

This case study illustrates our proposed theoretical model for eHealth behavior change interventions: simulating a counselor's support of working alliance, internalization of motivation, lapse preparation, and lapse management simultaneously. The case, *Endre*, is a fully automated smoking cessation program where each session takes the form of a written "counseling session" between the user and the program. The program content and structure were analyzed using intervention mapping [38],

illustrating the translation from theoretical model to intervention. The analysis shows that simulation of the 3 therapeutic processes is accomplished through a range of program elements. Working alliance [7,8] is supported through alliance factors [13], a nonembodied relational agent [12],cMI (Multimedia Appendix 4), and dynamic tailoring [42]. Internal motivation [9] is supported through cMI, dynamic tailoring, and modeling [38]. Finally, relapse is sought prevented through a psychoeducative session on lapses and relapses and a postquit day lapse management component.

By defining the components of a program and discussing its potentials for interactivity and tailoring in terms of concepts from face-to-face counseling, eHealth programs can be better understood [6]. This has implications both for program development and for the theoretical development of eHealth therapeutic process. In addition, by showing how the therapeutic processes of a program can be documented, from abstract concept through operationalization to simulation in specific program elements, we have demonstrated how intervention mapping used in a focused manner provides a compelling, interpretative approach to eHealth case studies. The value of such an inquiry for future empirical investigation is substantial: If the intervention should prove not to be effective, this may be because the identified theoretical processes are not sufficient for supporting behavior change or because the translation from theory to intervention elements was less than optimal.

The analysis of *Endre* suggests that the simultaneous simulation of each therapeutic process may result in a synergy effect. The operationalization in Table 1 reveals some of these potential interaction effects. The table visualizes that a working alliance is also a part of internalized motivation. When a working alliance to *Endre* is supported, this can influence the user's need for relatedness, thus supporting his or her internalized motivation to quit [27]. In addition, Table 1 visualizes that a working alliance and internalized motivation (columns) cross behavior maintenance (row). This means that for *Endre* to succeed in helping the user manage lapses, he or she needs to have both a working alliance to *Endre* and internalized motivation to recover

from a lapse, demonstrating that lapse management in a fully automated program can benefit from a strong working alliance and internalized motivation. A strong working alliance may enhance the effect of a lapse management program element through facilitating client processes such as commitment, satisfaction, and trust [23].This may increase the likelihood of the user staying with the program long enough to benefit from the lapse management therapy and trust the therapy that is given. At the same time, internalized motivation increases self-regulation, performance, and persistence [9,27,28] and may function as a buffer for future lapses. Should the user experience a lapse, a program that is supportive through that difficult period is likely to strengthen the working alliance by demonstrating sensitivity to the user's changing needs [23]. Furthermore, if the user should succeed in overcoming the lapse it would also presumably increase his or her feeling of competence, again enhancing internalized motivation [27]. It seems therefore that simultaneous simulation of supporting a working alliance, internalized motivation, and lapse management may result in a mutual enhancement of each process. These hypothesized synergy effects are displayed in Figure 10.

Interaction can be assumed from the operationalization level, but the step to simulation also shows the many methods and program elements that support several therapeutic processes at once. For example, all program material is delivered by the relational agent *Endre* using cMI. A relational agent supports working alliance [12], and cMI supports both working alliance [25] and internalized motivation [26], but in different ways. *Endre* also uses cMI in the lapse management session, influencing all 3 therapeutic processes at once. Another example of a program element that support several therapeutic processes are the 4 "quitters," serving both as guides in how to do the program exercises (supporting a working alliance) and as models in how to exercise autonomy in the quitting process (supporting internalized motivation). The fact that many program elements support several therapeutic processes at once implies that the effort needed to incorporate more than one therapeutic process in a program may diminish for each process included.

Figure 10. Hypothesized synergy effects of the 3 therapeutic processes.
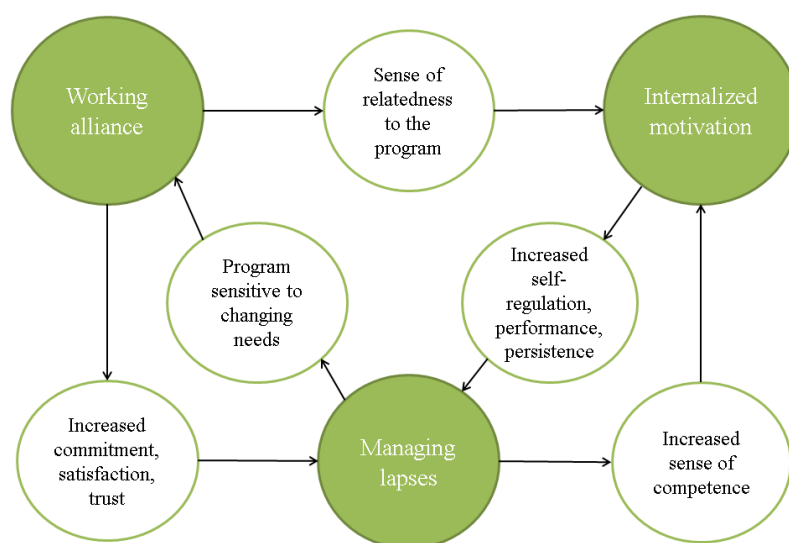
Comparison with Prior Work

All 3 therapeutic processes have been applied to fully automated programs previously. Studies on working alliance have mostly been on relational agents [11,12]. *Endre* builds on this work, although applying a nonembodied, rather than embodied relational agent, allowing the user freedom to "create" aspects of the relational agent. To further support a working alliance, *Endre* also incorporates alliance factors [13]. "Endre" also builds on previous work in the application of MI [32,33] and use of lapse preparation and lapse management [31,34-37]. The most significant contribution of *Endre*, however, is simulating all 3 processes simultaneously, something that to the best of our knowledge has not been done before systematically in a fully automated eHealth program.

Finally, this paper extends earlier work using intervention mapping eHealth tools to present a focused descriptive analysis of chosen program elements. Papers that use intervention mapping usually follow the structure of the intervention mapping steps and reports on most of these [29,30,40-50]. Instead of giving a full account of the breadth of the program, this paper uses intervention mapping for a focused descriptive analysis to make an argument of possible important eHealth elements. The description is intended to be sufficiently deep to allow for further inquiry into the chosen elements. This application of intervention mapping represents a complementary approach to the standard use of the method, that is, instead of using intervention mapping as a purely descriptive tool, we use it as a normative tool to determine what elements *should* be present in the "black box" of eHealth programs.

Limitations

Although comprehensive, the analysis presented here is a simplification of how the 3 therapeutic processes are simulated in *Endre*. Especially the social behavior of the relational agent, cMI, and dynamic tailoring are elements that are used in the entire program, and a full account was therefore not possible. Another limitation is that to highlight the 3 therapeutic

processes, descriptive depth was chosen over descriptive breadth. In addition, *Endre* does not simulate the 3 therapeutic processes perfectly. A fully automated program neither has the flexibility nor the presence of an actual human being. Just as *Endre* is not a human counselor, cMI is not MI. But the program may nevertheless simulate these 3 therapeutic processes convincingly enough to derive some of the benefits they have in face-to-face counseling. It should also be noted that *Endre* only represents one way in which these therapeutic processes may be simulated. Thus, if *Endre* fails to be an efficient program, it may be because the therapeutic processes in a fully automated program are not successful in inducing change or because the simulation of the therapeutic processes in *Endre* was inadequate.

There are, of course, limitations to the type of program that *Endre* represents. First, not everyone who wishes to quit smoking may benefit from such a detailed program. In the first author's clinical experience, some simply quit and do not wish to spend more time elaborating on the process. A participant in an earlier study conducted by the third author [57] actually experienced late night SMS messages asking whether she had been smoke free that day as smoking cues, creating a risk of (re)lapsing. *Endre* does make it possible for "unproblematic" quitters to move through many of the sessions rapidly, and the flexible session manager makes it possible to complete fewer sessions than what is in the full program. Nevertheless, it is a quite extensive intervention, communicating an expectation that quitting smoking is a process instead of a one-time action and requiring answers to daily SMS messages. Second, not everyone may wish to convey their thoughts with a program. Efforts to simulate a therapeutic setting aside, the therapy may still seem too artificial and ultimately unconvincing to the user. Alternately, the simulation may be too convincing, and sharing one's personal thoughts on quitting smoking with a machine that responds empathically to one's input may create an "uncanny" feeling [58] because the program acts like a human without being one. Even though reports from users of *Endre* so far indicate to the contrary that they respond positively to the

"mixture" of Man and machine, this is an area that will require further research.

Future Directions

Endre and the theoretical model presented here will be evaluated in forthcoming studies. Because the application of the model to the program is made explicit, it is possible to test. Empirical investigations may in turn influence or alter the theoretical model or its recommended application to a program.

In one current Randomized Controlled Trial (RCT), the lapse management component will be evaluated by randomly allocating participants to one version of the program with the lapse management component and one version without it. The results of this RCT will tell us whether providing immediate help to users who have had a lapse can significantly improve their success rate. We also plan to collect indicators on working alliance and on internal motivation.

Another ongoing project is a qualitative study on the users' working alliance to *Endre*. The goal of this study is to explore the nature of the working alliance because it is not given that working alliance to a fully automated program is identical to the working alliance to a human therapist. It is only when we can be convinced of the nature of working alliance to a fully automated program that it will be truly meaningful to test its importance for eHealth-assisted behavior change.

Finally, although we have argued that *Endre* simulates support of a working alliance, internalized motivation, and lapse preparation and lapse management, we do not know to what extent this simulation is successful for the user. One might establish simulation success through RCTs as the one described previously and compare the results with comparable findings from the counseling literature.

Conclusions

We have demonstrated how *Endre*, a fully automated eHealth program, through interactivity and individual tailoring emulate 3 effective mechanisms of face-to-face counseling. By having used intervention mapping to systematically break down *Endre* into smaller components and showed how the program simulates a counselor's support of a working alliance, internalized motivation, and lapse preparation and lapse management, our analysis is an example of how knowledge of what works in eHealth programs may be deepened by interpreting them in light of therapeutic processes. We suggest that the combination of these 3 therapeutic processes may result in a synergistic effect. Based on the analysis, we believe the combined support of a working alliance, internalization of motivation, and lapse preparation and management should be an element in the "black box" of automated eHealth behavior change programs that will make them more effective.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Overview of the themes in Endre.

[PDF File (Adobe PDF File), 366KB - jmir_v18i6e176_app1.pdf]

Multimedia Appendix 2

The postquit day lapse management component of Endre.

[PDF File (Adobe PDF File), 345KB - jmir_v18i6e176_app2.pdf]

Multimedia Appendix 3

Change model from the development of Endre.

[PDF File (Adobe PDF File), 216KB - jmir_v18i6e176_app3.pdf]

XSL•FO

RenderX

Multimedia Appendix 4

Computerized motivational interviewing in Endre.

[[PDF File (Adobe PDF File), 476KB](#) - [jmir_v18i6e176_app4.pdf](#)]

References

1. Kok G, Mesters I. Getting inside the black box of health promotion programmes using intervention Mapping. Chronic Illn 2011 Sep;7(3):176-180. [doi: 10.1177/1742395311403013] [Medline: 21900338]

2. Michie S, Abraham C, Eccles MP, Francis JJ, Hardeman W, Johnston M. Strengthening evaluation and implementation by specifying components of behaviour change interventions: a study protocol. Implement Sci 2011;6:10 [FREE Full text] [doi: 10.1186/1748-5908-6-10] [Medline: 21299860]

3. Dombrowski S, Sniehotta F, Avenell A. Current issues and future directions in Psychology and Health: Towards a cumulative science of behaviour change: Do current conduct and reporting of behavioural interventions fall short of best practice? Psychology & Health 2007 Dec;22(8):869-874. [doi: 10.1080/08870440701520973]

4. Strecher V. The Internet: Just another smoking cessation tool? Addiction 2008 Mar;103(3):485-486. [doi: 10.1111/j.1360-0443.2008.02144.x]

5. 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]

6. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? Transl Behav Med 2011 Mar;1(1):53-71 [FREE Full text] [doi: 10.1007/s13142-011-0021-7] [Medline: 21796270]

7. Safran J, Muran J. Negotiating the Therapeutic Alliance : A Relational Treatment Guide. New York: The Guilford Press; 2003.

8. Wampold B. The great psychotherapy debate: models, methods, and findings. Mahwah, NJ: L. Erlbaum Associates; 2001.

9. Ryan RM, Deci EL. Overview of self-determination theory: An organismic dialectical perspective. In: The Handbook of Self-Determination Research. Rochester: University of Rochester Press; 2002.

10. Marlatt GA, George WH. Relapse prevention: introduction and overview of the model. Br J Addict 1984 Sep;79(3):261-273. [Medline: 6595020]

11. Bickmore TW, Picard RW. Establishing and maintaining long-term human-computer relationships. ACM Trans Comput-Hum Interact 2005 Jun 01;12(2):293-327. [doi: 10.1145/1067860.1067867]

12. 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]

13. Barazzone N, Cavanagh K, Richards DA. Computerized cognitive behavioural therapy and the therapeutic alliance: a qualitative enquiry. Br J Clin Psychol 2012 Nov;51(4):396-417. [doi: 10.1111/j.2044-8260.2012.02035.x] [Medline: 23078210]

14. Ormrod JA, Kennedy L, Scott J, Cavanagh K. Computerised cognitive behavioural therapy in an adult mental health service: a pilot study of outcomes and alliance. Cogn Behav Ther 2010;39(3):188-192. [doi: 10.1080/16506071003675614] [Medline: 20485996]

15. Peck DF. The therapist-client relationship, computerized self-help and active therapy ingredients. Clin Psychol Psychother 2010;17(2):147-153. [doi: 10.1002/cpp.669] [Medline: 19946947]

16. Meyer B, Bierbrodt J, Schröder J, Berger T, Beevers C, Weiss M, et al. Effects of an Internet intervention (Deprexis) on severe depression symptoms: Randomized controlled trial. Internet Interventions 2015 Mar;2(1):48-59. [doi: 10.1016/j.invent.2014.12.003]

17. Morris ME, Aguilera A. Mobile, Social, and Wearable Computing and the Evolution of Psychological Practice. Prof Psychol Res Pr 2012 Dec;43(6):622-626 [FREE Full text] [doi: 10.1037/a0029041] [Medline: 25587207]

18. Danaher BG, Brendryen H, Seeley JR, Tyler MS, Woolley T. From black box to toolbox: Outlining device functionality, engagement activities, and the pervasive information architecture of mHealth interventions. Internet Interv 2015 Mar 1;2(1):91-101. [doi: 10.1016/j.invent.2015.01.002] [Medline: 25750862]

19. De Boni M, Richardson A, Hurling R. Humour, Relationship Maintenance and Personality Matching in automated dialogue: A controlled study. Interacting with Computers 2008 May;20(3):342-353. [doi: 10.1016/j.intcom.2007.11.002]

20. Cavanagh K, Millings A. (Inter)personal Computing: The Role of the Therapeutic Relationship in E-mental Health. J Contemp Psychother 2013 Jul 17;43(4):197-206. [doi: 10.1007/s10879-013-9242-z]

21. Brandt CL, Dalum P, Thomsen TT. "I miss the care even though I know it's just a machine": an explorative study of the relationship between an Internet-based smoking cessation intervention and its participants. Health Informatics J 2013 Sep;19(3):233-243. [doi: 10.1177/1460458212470572] [Medline: 23981397]

22. Kaplan B, Farzanfar R, Friedman RH. Personal relationships with an intelligent interactive telephone health behavior advisor system: a multimethod study using surveys and ethnographic interviews. Int J Med Inform 2003 Aug;71(1):33-41. [Medline: 12909156]

23.   Cahill J, Barkham M, Hardy G, Gilbody S, Richards D, Bower P, et al. A review and critical appraisal of measures of therapist-patient interactions in mental health settings. Health Technol Assess 2008 Jun;12(24):iii, ix-iii, 47 [FREE Full text] [Medline: 18510875]

24.   Miller WR, Rollnick S. Motivational Interviewing, Third Edition: Helping People Change (Applications of Motivational Interviewing). New York: The Guilford Press; 2012.

25.   Lundahl BW, Kunz C, Brownell C, Tollefson D, Burke BL. A Meta-Analysis of Motivational Interviewing: Twenty-Five Years of Empirical Studies. Research on Social Work Practice 2010 Jan 11;20(2):137-160. [doi: 10.1177/1049731509347850]

26.   Markland D, Ryan RM, Tobin VJ, Rollnick S. Motivational Interviewing and Self–Determination Theory. Journal of Social and Clinical Psychology 2005 Sep;24(6):811-831. [doi: 10.1521/jscp.2005.24.6.811]

27.   Patrick H, Williams GC. Self-determination theory: its application to health behavior and complementarity with motivational interviewing. Int J Behav Nutr Phys Act 2012;9:18 [FREE Full text] [doi: 10.1186/1479-5868-9-18] [Medline: 22385676]

28.   Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [Medline: 11392867]

29.   Voogt CV, Poelen EA, Kleinjan M, Lemmers LA, Engels RC. The development of a web-based brief alcohol intervention in reducing heavy drinking among college students: an Intervention Mapping approach. Health Promot Int 2013 Dec;29(4):669-679 [FREE Full text] [doi: 10.1093/heapro/dat016] [Medline: 23525645]

30.   Dalum P, Schaalma H, Kok G. The development of an adolescent smoking cessation intervention--an Intervention Mapping approach to planning. Health Educ Res 2012 Feb;27(1):172-181 [FREE Full text] [doi: 10.1093/her/cyr044] [Medline: 21730251]

31.   Tensil M, Jonas B, Strüber E. Two fully automated web-based interventions for risky alcohol use: randomized controlled trial. J Med Internet Res 2013;15(6):e110 [FREE Full text] [doi: 10.2196/jmir.2489] [Medline: 23742808]

32.   Friederichs SA, Oenema A, Bolman C, Guyaux J, van Keulen H, Lechner L. I Move: systematic development of a web-based computer tailored physical activity intervention, based on motivational interviewing and self-determination theory. BMC Public Health 2014;14:212 [FREE Full text] [doi: 10.1186/1471-2458-14-212] [Medline: 24580802]

33.   Moreau M, Gagnon M, Boudreau F. Development of a fully automated, web-based, tailored intervention promoting regular physical activity among insufficiently active adults with type 2 diabetes: integrating the I-change model, self-determination theory, and motivational interviewing components. JMIR Res Protoc 2015;4(1):e25 [FREE Full text] [doi: 10.2196/resprot.4099] [Medline: 25691346]

34.   Brendryen H, Kraft P, Schaalma H. Looking Inside the Black Box: Using Intervention Mapping to Describe the Development of the Automated Smoking Cessation Intervention 'Happy Ending'. The Journal of Smoking Cessation 2010 Jun;5(1):29-56. [doi: 10.1375/jsc.5.1.29]

35.   Brendryen H, Johansen A, Nesvåg S, Kok G, Duckert F. Constructing a Theory- and Evidence-Based Treatment Rationale for Complex eHealth Interventions: Development of an Online Alcohol Intervention Using an Intervention Mapping Approach. JMIR Res Protoc 2013;2(1):e6 [FREE Full text] [doi: 10.2196/resprot.2371] [Medline: 23612478]

36.   Rose GL, Skelly JM, Badger GJ, Naylor MR, Helzer JE. Interactive voice response for relapse prevention following cognitive-behavioral therapy for alcohol use disorders: a pilot study. Psychol Serv 2012 May;9(2):174-184 [FREE Full text] [doi: 10.1037/a0027606] [Medline: 22662731]

37.   Elfeddali I, Bolman C, Candel MJ, Wiers RW, de Vries H. Preventing smoking relapse via Web-based computer-tailored feedback: a randomized controlled trial. J Med Internet Res 2012;14(4):e109 [FREE Full text] [doi: 10.2196/jmir.2057] [Medline: 22903145]

38.   Bartholomew L, Parcel G, Kok G, Gottlieb N, Fernandez M. Planning Health Promotion Programs: An Intervention Mapping Approach. San Fransisco, CA: Jossey-Bass; 2011.

39.   Danaher BG, McKay HG, Seeley JR. The information architecture of behavior change websites. J Med Internet Res 2005;7(2):e12 [FREE Full text] [doi: 10.2196/jmir.7.2.e12] [Medline: 15914459]

40.   Sassen B, Kok G, Mesters I, Crutzen R, Cremers A, Vanhees L. A web-based intervention for health professionals and patients to decrease cardiovascular risk attributable to physical inactivity: development process. JMIR Res Protoc 2012;1(2):e21 [FREE Full text] [doi: 10.2196/resprot.1804] [Medline: 23612470]

41.   Kok G, Harterink P, Vriens P, Zwart O, Hospers HJ. The gay cruise: Developing a theory- and evidence-based Internet HIV-prevention intervention. Sex Res Soc Policy 2006 Jun;3(2):52-67. [doi: 10.1525/srsp.2006.3.2.52]

42.   Jacobs NC, Völlink T, Dehue F, Lechner L. Online Pestkoppenstoppen: systematic and theory-based development of a web-based tailored intervention for adolescent cyberbully victims to combat and prevent cyberbullying. BMC Public Health 2014;14:396 [FREE Full text] [doi: 10.1186/1471-2458-14-396] [Medline: 24758264]

43.   Vonk NA, Huirne JA, Pittens CA, van Mechelen WR, Broerse JE, Brölmann HA, et al. eHealth program to empower patients in returning to normal activities and work after gynecological surgery: intervention mapping as a useful method for development. J Med Internet Res 2012;14(5):e124 [FREE Full text] [doi: 10.2196/jmir.1915] [Medline: 23086834]

44.   van Genugten L, van Empelen P, Flink I, Oenema A. Systematic development of a self-regulation weight-management intervention for overweight adults. BMC Public Health 2010;10:649 [FREE Full text] [doi: 10.1186/1471-2458-10-649] [Medline: 20979603]

45. Robben SH, Huisjes M, van Achterberg T, Zuidema SU, Olde Rikkert MG, Schers HJ, ZOWEL NN Study Group. Filling the Gaps in a Fragmented Health Care System: Development of the Health and Welfare Information Portal (ZWIP). JMIR Res Protoc 2012;1(2):e10 [FREE Full text] [doi: 10.2196/resprot.1945] [Medline: 23611877]
46. Heinrich E, Schaper N, de Vries N. Development of the web-based type 2 diabetes education programme: DIEP. European Diabetes Nursing 2009 Sep;6(2):51-56. [doi: 10.1002/edn.135]
47. Drozd F, Haga SM, Brendryen H, Slinning K. An Internet-Based Intervention (Mamma Mia) for Postpartum Depression: Mapping the Development from Theory to Practice. JMIR Res Protoc 2015;4(4):e120 [FREE Full text] [doi: 10.2196/resprot.4858] [Medline: 26476481]
48. Plaete J, De Bourdeaudhuij BI, Verloigne M, Oenema A, Crombez G. A Self-Regulation eHealth Intervention to Increase Healthy Behavior Through General Practice: Protocol and Systematic Development. JMIR Res Protoc 2015;4(4):e141 [FREE Full text] [doi: 10.2196/resprot.4835] [Medline: 26694101]
49. Oosterom-Calo R, Te Velde SJ, Stut W, Brug J. Development of Motivate4Change Using the Intervention Mapping Protocol: An Interactive Technology Physical Activity and Medication Adherence Promotion Program for Hospitalized Heart Failure Patients. JMIR Res Protoc 2015;4(3):e88 [FREE Full text] [doi: 10.2196/resprot.4282] [Medline: 26195072]
50. Geense WW, van Gaal BG, Knoll JL, Cornelissen EA, Schoonhoven L, Kok G. Online Support Program for Parents of Children With a Chronic Kidney Disease Using Intervention Mapping: A Development and Evaluation Protocol. JMIR Res Protoc 2016;5(1):e1 [FREE Full text] [doi: 10.2196/resprot.4837] [Medline: 26764218]
51. Park EW, Tudiver FG, Campbell T. Enhancing partner support to improve smoking cessation. Cochrane Database Syst Rev 2012;7:CD002928. [doi: 10.1002/14651858.CD002928.pub3] [Medline: 22786483]
52. Prescott P, Børtveit T. Helse og atferdsendring. Oslo: Gyldendal akademisk; 2004.
53. Batra A, Collins SE, Torchalla I, Schröter M, Buchkremer G. Multidimensional smoker profiles and their prediction of smoking following a pharmacobehavioral intervention. J Subst Abuse Treat 2008 Jul;35(1):41-52. [doi: 10.1016/j.jsat.2007.08.006] [Medline: 17931825]
54. Carver C, Scheier M. On the Self-Regulation of Behavior. Cambridge: Cambridge University Press; 1998.
55. Marlatt GA, Donovan DM, editors. Relapse Prevention: Maintenance Strategies in the Treatment of Addictive Behaviors. 2nd edition. New York: The Guilford Press; 2005.
56. Beck AT, Haigh EA. Advances in cognitive theory and therapy: the generic cognitive model. Annu Rev Clin Psychol 2014;10:1-24. [doi: 10.1146/annurev-clinpsy-032813-153734] [Medline: 24387236]
57. 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]
58. Mori M, MacDorman K, Kageki N. The Uncanny Valley [From the Field]. IEEE Robot Automat Mag 2012 Jun;19(2):98-100. [doi: 10.1109/MRA.2012.2192811]

Abbreviations

CMI: computerized motivational interviewing
MI: motivational interviewing
RCT: randomized controlled trial
SMS: short message service

XSL•FO
RenderX

Original Paper

Adherence to Internet-Based Mobile-Supported Stress Management: A Pooled Analysis of Individual Participant Data From Three Randomized Controlled Trials

Anna-Carlotta Zarski[1,2], MSc; Dirk Lehr[2,3], PhD; Matthias Berking[1,2], PhD; Heleen Riper[2,4,5], PhD; Pim Cuijpers[2,4], PhD; David Daniel Ebert[1,2], PhD

[1]Department of Clinical Psychology and Psychotherapy, Friedrich-Alexander University Erlangen-Nuremberg, Erlangen, Germany

[2]Innovation Incubator, Division of Online Health Training, Leuphana University Lueneburg, Luenburg, Germany

[3]Institute for Psychology, Department of Health Psychology and Applied Biological Psychology, Leuphana University Luenburg, Lueneburg, Germany

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

[5]Telepsychiatric Centre, University of Southern Denmark, Odense, Denmark

Corresponding Author:
Anna-Carlotta Zarski, MSc
Department of Clinical Psychology and Psychotherapy
Friedrich-Alexander University Erlangen-Nuremberg
Nägelsbachstr. 25a
Erlangen, 91052
Germany
Phone: 49 9131 85 67570
Fax: 49 4131 677 171
Email: Anna-Carlotta.Zarski@fau.de

Abstract

Background: Nonadherence to treatment is a prevalent issue in Internet interventions. Guidance from health care professionals has been found to increase treatment adherence rates in Internet interventions for a range of physical and mental disorders. Evaluating different guidance formats of varying intensity is important, particularly with respect to improvement of effectiveness and cost-effectiveness. Identifying predictors of nonadherence allows for the opportunity to better adapt Internet interventions to the needs of participants especially at risk for discontinuing treatment.

Objective: The goal of this study was to investigate the influence of different guidance formats (content-focused guidance, adherence-focused guidance, and administrative guidance) on adherence and to identify predictors of nonadherence in an Internet-based mobile-supported stress management intervention (ie, GET.ON Stress) for employees.

Methods: The data from the groups who received the intervention were pooled from three randomized controlled trials (RCTs) that evaluated the efficacy of the same Internet-based mobile-supported stress management intervention (N=395). The RCTs only differed in terms of the guidance format (content-focused guidance vs waitlist control, adherence-focused guidance vs waitlist control, administrative guidance vs waitlist control). Adherence was defined by the number of completed treatment modules (0-7). An ANOVA was performed to compare the adherence rates from the different guidance formats. Multiple hierarchical linear regression analysis was conducted to evaluate predictors of nonadherence, which included gender, age, education, symptom-related factors, and hope for improvement.

Results: In all, 70.5% (93/132) of the content-focused guidance sample, 68.9% (91/132) of the adherence-focused guidance sample, and 42.0% (55/131) of the participants in the administrative guidance sample completed all treatment modules. Guidance had a significant effect on treatment adherence ($F_{2,392}$=11.64, P<.001; ω^2=.05). Participants in the content-focused guidance (mean 5.70, SD 2.32) and adherence-focused guidance samples (mean 5.58, SD 2.33) completed significantly more modules than participants in the administrative guidance sample (mean 4.36, SD 2.78; t_{223}=4.53, P<.001; r=.29). Content-focused guidance was not significantly associated with higher adherence compared to adherence-focused guidance (t_{262}=0.42, P=.67; r=.03). The effect size of r=.03 (95% CI −0.09 to 0.15) did not pass the equivalence margin of r=.20 and the upper bound of the 95% CI lay below the predefined margin, indicating equivalence between adherence-focused guidance and content-focused guidance. Beyond the influence of guidance, none of the predictors significantly predicted nonadherence.

XSL•FO
RenderX

Conclusions: Guidance has been shown to be an influential factor in promoting adherence to an Internet-based mobile-supported stress management intervention. Adherence-focused guidance, which included email reminders and feedback on demand, was equivalent to content-focused guidance with regular feedback while requiring only approximately a quarter of the coaching resources. This could be a promising discovery in terms of cost-effectiveness. However, even after considering guidance, sociodemographic, and symptom-related characteristics, most interindividual differences in nonadherence remain unexplained.

Clinical Trial: DRKS00004749; http://drks-neu.uniklinik-freiburg.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL _ID=DRKS00004749 (Archived by WebCite at http://www.webcitation.org/6QiDk9Zn8); DRKS00005112; http://drks-neu.uniklinik-freiburg. de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00005112 (Archived by WebCite at http://www.webcitation.org/6QiDysvev); DRKS00005384; http://drks-neu.uniklinik-freiburg.de/ drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00005384 (Archived by WebCite at http://www.webcitation.org/6QiE0xcpE)

Introduction

Many participants of Internet-based programs do not begin the interventions after registration or are nonadherent and quit the intervention prematurely against recommendations in both study and routine care settings [1-6]. Treatment adherence can be defined as the extent to which individuals experience the intervention content [7]. Low treatment adherence is a major concern because it has been associated with reduced efficacy of Internet interventions for physical and mental disorders and health promotion programs [8-13].

Occupational stress is associated with an increased risk for common mental disorders in the long term [14]. Internet-based stress management interventions for employees show promising effects in reducing stress and related health problems such as depression [15-19]. Nevertheless, low treatment adherence is a serious problem that adversely impacts the efficacy of Internet-based antistress programs [20-22].

In order to improve the clinical effects of Internet interventions, it is important to identify and evaluate factors associated with adherence [23]. Treatment factors, in particular human guidance, may be especially important [24-26]. Guided treatments have been shown to result in higher adherence rates (56% to 81%) than unguided treatments (26% to 69%) [27-31]. Treatment adherence rates of guided interventions have often been found to be comparable to those of face-to-face interventions [32,33]. However, guidance from a health care professional is expensive and a limited resource. Thus, it is crucial to investigate the effects of different guidance formats to identify the level and type of guidance necessary to achieve acceptable treatment adherence.

Guidance in Internet interventions is often classified according to the amount of required coaching or therapist time [34]. However, guidance also differs regarding objectives and content. Accordingly, at least four different guidance formats can be distinguished: (1) unguided interventions, completely self-administered by the user; (2) administrative guidance, providing technical support in case of computer and Internet platform-related problems and dispensing relevant information throughout the course of the treatment; (3) adherence-focused guidance, with adherence monitoring including reminders by email or telephone; and (4) content-focused guidance, personalized written feedback by coaches for completed treatment modules and adherence monitoring.

Content-focused guidance in Internet interventions has been found to be associated with higher levels of treatment completion (72%) compared to administrative guidance (62%) or unguided interventions (26%) in a meta-analysis [31]. Adherence-focused guidance also seems promising in fostering treatment completion with the added benefit of keeping coaching demands—in terms of time spent per participant—to a minimum [26,35-39].

However, whether adherence-focused guidance results in comparable adherence rates to content-focused guidance formats remains unclear. Comparisons of adherence-focused guidance to content-focused guidance are currently limited regarding treatment adherence in Internet interventions.

Mohr and colleagues [25] introduced the notion of "supportive accountability" to explain the relationship between human guidance and treatment adherence in Internet interventions. This model assumes that human support in Internet interventions enhances adherence by allowing the patient to foster a commitment toward an eCoach, who is perceived as being trustworthy, benevolent, and professionally knowledgeable. The eCoach is responsible for accompanying the participants through the program, showing interest in their training processes, and offering support. In this respect, adherence monitoring in addition to reminders in case of nonadherence must be embedded in a benevolent context, with the aim of supporting adherence instead of surveillance of module completion. The expectations placed on the participant by the therapy program, to be sustained on a regular basis throughout the treatment, should be transparent and reasonable. Hence, the participant knows what is to be expected and can be involved in determining these expectations. Both in adherence- and content-focused guidance, the eCoach communicates the requirements for successful participation in the intervention and the associated expectations placed on the participants. By contrast, in unguided or administrative-guided interventions, participants only receive general recommendations on program use.

XSL•FO
RenderX

Based on the supportive accountability model [25], we assumed that the introduction of an eCoach who offers support in program completion is important for creating an adherence-promoting relationship with the participant. For this reason, we developed the concept of adherence-focused guidance, which is comprised of not only adherence monitoring, but also the opportunity to receive content feedback on demand [40]. At the beginning of the training, the eCoach invites the participant to contact them in case of content-related questions and any desired feedback for completed treatment modules. Feedback on demand is expected to emphasize the supporting role of the eCoach in the training process and thereby promote the participant's adherence. Assuming that accountability is the essential factor in coaching that keeps participants involved in training, we hypothesized that both adherence-focused guidance and content-focused guidance are superior to administrative guidance with regard to adherence rates. Adherence-focused guidance was expected to be equivalent to content-focused guidance for adherence rates.

Apart from treatment factors, knowledge of user characteristics related to nonadherence helps to identify individuals who are at risk of discontinuing treatment and might need additional support. According to the behavior change model for Internet interventions by Ritterband and colleagues [41], user characteristics can be classified into disease-related factors, demographics, traits, cognitive factors, beliefs and attitudes, physiological factors, and skills.

Demographic and disease-specific user characteristics were of particular interest in past adherence research in the field of Internet interventions. Low education level [1,42-44], male gender [1,30,42,45], or being unmarried or single [43,46-48] have been found to be associated with lower treatment adherence across different mental and physical health interventions. Younger age has also been shown to be related to nonadherence in the majority of studies [37,42,45,49]. Only a small number of studies found older individuals to be at a higher risk for nonadherence [1,2,7]. However, some studies found no significant association between any sociodemographic variables and treatment adherence [7,50,51].

High symptom severity at baseline is also frequently linked to lower treatment adherence [7,47,52,53]. But, the relationship between baseline depressive symptoms and treatment adherence seems inconsistent. Both higher [7,54,55] and lower [1,3,42] baseline depression scores were found to be associated with lower treatment adherence or no significant association was found [30,50,56,57]. Lower treatment expectations have been found to be related to nonadherence in Internet interventions [56,58,59].

Only a few studies to date have investigated predictors of adherence in Internet intervention and they showed conflicting results [60]. For this reason, we used an exploratory approach including potential predictors based on (1) results of previous studies investigating adherence predictors in Internet interventions [1,7,42-44] and (2) theoretical assumptions due to intervention characteristics. The final list of potential predictors investigated in the present study included: (1) demographics: gender, education level, and age; (2) symptom-related factors: stress, depression, and emotional exhaustion; and (3) variables concerning beliefs and attitudes: hope for improvement.

To the best of our knowledge, no research to date has explored the influence of different guidance formats and user characteristics on treatment adherence to an Internet intervention in stressed employees.

The current study aimed to (1) report adherence rates from a newly developed Internet-based mobile-supported stress management intervention (ie, GET.ON Stress) and (2) investigate the role of different guidance formats (content-focused guidance, adherence-focused guidance, and administrative guidance) on adherence. Further goals of the study were to (3) identify user characteristics predictive of treatment nonadherence over and above the guidance formats and (4) analyze differential predictor effects as a function of guidance formats.

We hypothesized that (1) treatment adherence rates would be greater for content-focused guidance and adherence-focused guidance compared to administrative guidance, and (2) adherence-focused guidance is equivalent to content-focused guidance in terms of adherence rates. User characteristics that contribute to treatment nonadherence apart from guidance formats and differential effects of predictors as a function of guidance formats were analyzed exploratively. In this study, treatment adherence is operationalized by the number of completed treatment modules.

Methods

Data for this analysis were drawn from three randomized controlled trials (RCTs) evaluating the same Internet-based mobile-supported stress management intervention (GET.ON Stress [61-64]) under varying guidance conditions (study 1: content-focused guidance vs waitlist control; study 2: adherence-focused guidance vs waitlist control; study 3: administrative guidance vs waitlist control). Details of the study design for study 1 have been described in a published study protocol [61]. All three trials employed the same design and procedures apart from the guidance format, allowing for the pooling of the data from the three studies.

Sample

The analyses in this study were based solely on the intervention group samples of the N=395 participants who received the same Internet-based mobile-supported stress management intervention (study 1: n=132; study 2: n=132; study 3: n=131). Participants in the waitlist control condition were not included in the analyses because they did not receive access to the training until 6 months after randomization. All three studies included (1) currently employed workers, (2) older than age 18 years, (3) with scores ≥22 on the Perceived Stress Scale (PSS-10) [65], (4) who self-reported having Internet access, (5) sufficient skills in reading and writing German, and (6) who were willing to give informed consent. Participants were excluded when (1) they self-reported having been diagnosed with psychoses or dissociative symptoms in the past or (2) showed a notable suicidal risk as indicated by a score of greater than 1 on item 9

XSL•FO

RenderX

("I feel I would be better off dead") on the Beck Depression Inventory (BDI) [66].

Intervention

The Internet-based mobile-supported stress management intervention GET.ON Stress is based on two main components: problem solving [16,67-69] and emotion regulation [70,71]. A detailed description can be found elsewhere [61]. The intervention consists of seven modules composed of psychoeducation (module 1), problem solving (modules 2-3), emotion regulation (modules 4-6), planning for the future (module 7), and an optional booster session. Additionally, participants are offered eight units that are integrated in modules 2 to 6 that can be opted for based on individual needs or preference. These units are directed at time management, rumination and worrying, psychological detachment from work, sleep hygiene, rhythm and regularity of sleeping habits, nutrition and exercise, organization of breaks during work, and social support. Each module can be completed in approximately 45 to 60 minutes. Participants were advised to do at least one and a maximum of two modules per week. Consequently, the intervention took approximately 4 to 7 weeks (not including the booster session offered 4 weeks after completion of the intervention). Lessons were in the format of text, exercises, and testimonials, and included interactive elements such as audio and video clips. Participants were encouraged to keep a daily online stress diary. One strong focus of the intervention lay in transfer tasks (homework assignments) to integrate the newly acquired strategies and techniques into daily life. The training was adaptive because the content is tailored to the specific needs of the individual participants by continuously asking them to choose among various response options. Using responsive Web design, participants could use the program on a computer, tablet, or mobile phone. An integrated read-aloud function allowed participants to follow narrated lessons. If desired, participants could receive automatic motivational text messages and small exercises on their mobile phones. These messages had the purpose of supporting the participant in transferring the exercises of the training into their daily lives (eg, short relaxation exercises). The participants had the opportunity to choose between "light coach" (one text message every other day) or "intensive coach" (two to three text messages every day) options.

Content-Focused Guided Internet-Based Mobile-Supported Stress Management Intervention

Participants in the content-focused guidance condition received personalized written feedback from an eCoach on the exercises they had completed in each module within 48 hours. The eCoaches were psychologists and trained Master's-level psychology students who followed guidelines about the feedback process that were defined according to the standardized manual for the intervention. The eCoaches were advised to not spend more than 30 minutes on feedback on a given completed module. The eCoaches sent reminders when the participants did not complete a module within 7 days. In total, the eCoaches sent 365 reminders, corresponding to a mean 2.77 reminders per participant (range 0-11, SD 2.41). The time required for coaching totaled up to 4 hours per participant.

Adherence-Focused Guided Internet-Based Mobile-Supported Stress Management Intervention

Participants of the adherence-focused guidance condition were also supported by an eCoach. The guidance manual was based on our developed adherence-focused guidance concept as outlined in the Introduction [40]. The eCoaches were trained psychologists who followed guidelines about the feedback process that were defined according to the standardized manual for the intervention. Adherence-focused guidance consisted of adherence monitoring and feedback on demand. Adherence monitoring included regularly checking module completion and sending reminders in case the participant did not complete at least one module within 7 days. In total, the eCoaches sent 463 reminders, corresponding to a mean 3.51 reminders per participant (range 0-13, SD 1.98). Feedback on demand included offering participants the opportunity to contact the eCoach and receive individual support or feedback within 24 hours. In contrast to the content-focused guidance concept, eCoach guidance only took place at the initiative of the participants to minimize the costs. There were only requests for a mean 8 content feedbacks by all participants (range 0-5, SD 0.46), corresponding to 0.06 feedback demands per participant. Thus, most of the time spent per participant was related to checking adherence to the intervention and providing reminders in case of nonadherence. The time required for coaching, including all reminders and feedback by request, added up to 1 hour per participant.

Administrative-Guided Internet-Based Mobile-Supported Stress Management Intervention

Participants in the administrative guidance condition were provided with contact information for the study administration team during the study period, which addressed such things as the completion of questionnaires, but they were not supported by an eCoach. They were provided with an email address to use in case of any technical problems.

Measures

Adherence Measures

The number of completed treatment modules in the Internet-based mobile-supported stress management intervention, which ranged from 0 to 7, was the primary outcome measure in this study and was assessed by the system that provided the intervention. Module completion was defined by completion of the last page of a module. To arrive on the last page, participants were required to complete all the previous writing tasks. A module completion score of 0 could either mean that the participant did not start the intervention or did not finish the first module. Each module took approximately 45 to 60 minutes for completion.

Predictor Measures

The following variables, assessed at baseline before the start of the program, were evaluated as potential predictors of nonadherence: sociodemographic factors (gender [male/female], age [years], level of education [low, middle, high]), symptom severity factors (perceived stress, depressive symptoms,

emotional exhaustion), and hope for improvement (confidence in treatment efficacy).

Perceived stress at baseline relating to the past week was examined with the German version of the 10-item Perceived Stress Scale (PSS-10) [65,72]. This self-report instrument uses a 5-point Likert-type scale that ranges from 0="never" to 4="very often." A Cronbach alpha of .77 indicated acceptable internal consistency of the PSS-10 in this study.

Baseline depression symptom severity was measured with the German version of the Center for Epidemiological Studies Depression Scale (CES-D) [73,74]. This frequently used self-report instrument consists of 15 items that are answered on a 4-point Likert-type scale and refer to the previous week. Total scores range from 0 to 60. In this study, internal consistency was good (Cronbach alpha=.88).

To measure emotional exhaustion, the basic stress dimension of burnout, the German version of the Maslach Burnout Inventory was utilized (MBI-GS-D) [75,76]. This commonly used self-report instrument consists of five items and uses a 6-point Likert-type scale anchored by 1="never" and 6="very often." In this study, internal consistency was acceptable (Cronbach alpha=.79).

Hope for improvement (confidence in treatment efficacy) was measured using the homonymous subscale of the German Patient Questionnaire on Therapy Expectation and Evaluation (PATHEV) [77] adapted to the online training context. The items are rated on a 5-point Likert-type scale. The hope for improvement subscale showed acceptable internal consistency (Cronbach alpha=.79) in this study.

Analysis

An ANOVA was conducted to compare the treatment adherence rates between the three guidance formats with guidance as the independent variable and adherence as the dependent variable [78]. The power analysis revealed that with the given sample size (N=395), small effects ($\omega^2 = 0.03$) with a power (1-beta) of 80% and alpha=.05 would have been detected. Omega squared (ω^2) was used as the measure for effect size of the ANOVA overall effect with values of 0.01, 0.06, and 0.14 representing small, medium, and large effects, respectively [79].

In planned contrasts, the superiority of the content-focused guidance and adherence-focused guidance over administrative guidance as well as equivalence of adherence-focused guidance and content-focused guidance was assessed. The effect sizes for the planned comparisons were described by r with values of .1, .3, and .5 indicating small, medium, and large effects, respectively [80]. In order to investigate the hypothesis that adherence-focused guidance is equivalent to content-focused guidance, a confidence interval approach was used on the effect size of the difference between adherence-focused guidance and content-focused guidance with a two-sided .05 level of significance [81]. The equivalence margin was set a priori at $r=.20$ corresponding to the smallest value that would present a relevant effect [82]. The upper bound of the 95% CI for the effect size was compared with the predefined equivalence margin of $r=.20$ and had to be below the margin to show equivalence, with a significance level of .05. Hierarchical multiple linear regression analysis was performed on the combined sample to explore potential predictors of nonadherence in addition to guidance formats with gender, age, education level, stress level, depression, emotional exhaustion, and hope for improvement as the independent variables, and adherence as the dependent variable [78]. In this model, sociodemographic variables (gender, age, and education level) were entered in the first step, followed by baseline symptom-related factors (stress level, depression, and emotional exhaustion) and hope for improvement in the second step, and guidance (content-focused guidance vs administrative guidance, adherence-focused guidance vs administrative guidance) in the third step. In the final step, the interactions between the guidance formats and the other predictor variables were added to the model to explore potentially differential predictor effects as a factor of guidance formats.

All continuous predictors were group mean centered. The power analysis revealed that with the given sample size (N=395), small effects ($r=.20$) with a power (1–beta) of 80% and Cronbach alpha=.05 would have been detected.

One participant had missing values in the depression level and the hope of improvement variable for which data were imputed using a Markov chain Monte Carlo multivariate imputation algorithm (missing data module in SPSS version 22) with 100 estimations per missing value.

In order to test the robustness of our results, we applied sensitivity analyses. We used (1) a more conservative outcome by defining modules as completed only when finished within 12 weeks and (2) Kaplan-Meier survival curves to compare the adherence rates between the three different guidance formats. All analyses were performed using SPSS version 22. Directed hypotheses were tested with a one-tailed test and nondirected hypotheses with a two-tailed test.

Results

Descriptive Statistics

In total, 395 participants were included in the analysis. Baseline characteristics of the study population are presented in Table 1. There was a significant difference in the gender ratio and the education level between the three studies. The adherence-focused guidance sample showed a significantly lower percentage of male participants (13.6%, 18/132) compared to the administrative guidance (26.0%, 34/131) and content-focused guidance samples (26.5%, 35/132). The content-focused guidance sample had a significantly higher education level (64.4%, 85/132) compared to the administrative guidance (56.5%, 74/131) and adherence-focused guidance samples (52.3%, 69/132). However, gender and education level were not associated with treatment adherence in the explorative analysis and, thus, not accounted for in subsequent analyses.

Table 1. Baseline characteristics of the study population (N=395).

Characteristic	Administrative guidance (n=131)	Adherence-focused guidance (n=132)	Content-focused guidance (n=132)	*P*
Age (years), mean (SD)	41.2 (9.4)	42.6 (9.5)	42.4 (10.7)	.51
Gender, n (%)				.02
Female	97 (74.1)	113 (85.6)	97 (73.5)	
Male	34 (26.0)	18 (13.6)	35 (26.5)	
Other		1 (0.8)		
Ethnicity, n (%)				.98
Caucasian	107 (81.7)	108 (81.8)	110 (83.3)	
Asian	1 (0.8)	1 (0.8)	0	
Not reported	23 (17.6)	23 (17.4)	22 (16.7)	
Marital status, n (%)				.96
Unmarried	40 (30.5)	39 (29.6)	43 (32.6)	
Married	65 (49.6)	62 (47.0)	63 (47.7)	
Cohabited	16 (12.2)	18 (13.6)	17 (12.9)	
Separated	9 (6.9)	13 (9.9)	8 (6.1)	
Widowed	1 (0.8)	0	1 (0.8)	
Education, n (%)				.01
Low	0	1 (0.8)	5 (3.8)	
Middle	57 (43.5)	62 (47.0)	42 (31.8)	
High	74 (56.5)	69 (52.3)	85 (64.4)	
Gross annual income (in Euro), n (%)				.34
Low	39 (29.8)	41 (31.1)	35 (26.5)	
Middle	40 (30.5)	33 (25.0)	26 (19.7)	
High	45 (34.4)	49 (37.1)	59 (44.7)	
Not reported	7 (5.3)	9 (6.8)	12 (9.1)	
Employment status, n (%)				.42
Permanent	104 (79.4)	107 (81.1)	110 (83.3)	
Temporary	19 (14.5)	14 (10.6)	11 (8.3)	
Self-employed	6 (4.6)	9 (6.8)	11 (8.3)	
Other	2 (1.5)	2 (1.5)	0	
Experience with health-related programs, n (%)				.82
Yes	14 (10.7)	17 (12.9)	17 (12.9)	
No	117 (89.3)	115 (87.1)	115 (87.1)	
Experience with face-to-face psychotherapy, n (%)				.32
Yes	47 (35.9)	46 (34.9)	57 (43.2)	
No	84 (64.1)	86 (65.2)	75 (56.8)	
Stress, mean (SD)	25.7 (5.0)	25.2 (4.6)	25.9 (3.9)	.44
Depression, mean (SD)	25.1[a] (9.3)	23.2 (9.3)	23.3 (8.5)	.15
Emotional exhaustion, mean (SD)	4.8 (0.8)	4.7 (0.8)	4.7 (0.7)	.96
Hope of improvement, mean (SD)	3.7[a] (0.6)	3.6 (0.6)	3.7 (0.7)	.27

[a] Due to missing data, the means refer to a subsample with n=130 in this group.

XSL•FO
RenderX

Adherence Rates Between Guidance Formats

Figures 1 and 2 depict the treatment adherence rates for all three samples. Figure 1 shows the total number of completed modules per participant and Figure 2 depicts the number of completed modules by module. In the administrative guidance sample, 13.7% (18/131) of the participants did not start the intervention, and 42.0% (55/131) completed all seven modules with a mean

of 4.4 completed modules (SD 2.8, range 0-7). In the group that received adherence-focused guidance, 5.3% (7/132) of the participants did not start the intervention and 68.9% (91/132) completed all modules with a mean number of 5.6 completed modules (SD 2.3, range 0-7). In the content-focused guidance sample, 7.6% (10/132) of the participants did not start the intervention and 70.5% (93/132) completed all modules with a mean number of 5.7 completed modules (SD 2.3, range 0-7).

Figure 1. Total number of completed modules per participant.
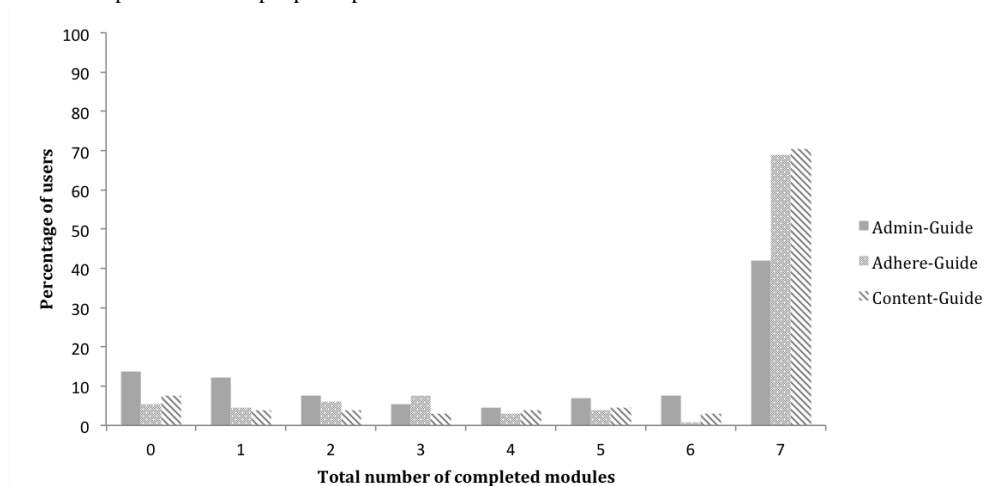
Figure 2. Number of completed modules by module.
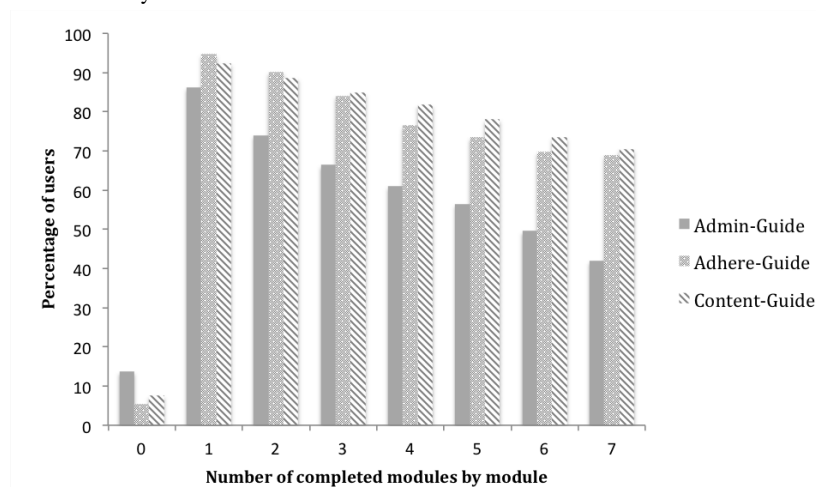
Guidance Formats on Adherence

As expected, there was a significant effect of guidance on treatment adherence ($F_{2,392}=11.64$, $P<.001$; $\omega^2=.05$). Planned comparisons revealed that (1) participants in the content-focused guidance (mean 5.7, SD 2.3) and adherence-focused guidance samples (mean 5.6, SD 2.3) completed significantly more modules than the participants in the administrative guidance sample (mean 4.4, SD 2.8; $t_{223}=4.53$, $P<.001$; $r=.29$) and (2) there was no significant difference between the content-focused guidance and adherence-focused guidance groups in treatment adherence ($t_{262}=0.42$, $P=.67$; $r=.03$). The effect size $r=.03$ (95% CI –0.09 to 0.15) did not cross the equivalence margin of $r=.20$. Likewise, the upper bound of the two-sided 95% confidence interval lay below the predefined margin, indicating that

adherence-focused guidance was equivalent to content-focused guidance. Hence, we conclude that adherence-focused guidance is not associated with relevant lower treatment adherence compared to content-focused guidance.

Using a more conservative outcome by defining modules as completed only when finished within 12 weeks did not result in any different conclusions (results not shown). Similarly, conducting survival analysis according to Kaplan-Meier, we derived comparable results. The survival distribution for administrative guidance was significantly different from the survival distribution of adherence-focused guidance ($\chi^2_2=19.0$, $P<.001$) and significantly different from the survival distribution of content-focused guidance ($\chi^2_2=21.7$, $P<.001$). No significant difference was found between the survival distributions of

adherence-focused guidance and content-focused guidance (χ^2_2=0.1, P=.77).

User Characteristics Predictive of Nonadherence

Table 2 shows the results of the imputed hierarchical multiple linear regression analysis examining predictors of treatment nonadherence over and above the guidance format. Due to missing values, the depression level variable and the hope of improvement variable of one participant had to be imputed. Entering the sociodemographic variables (gender, age, and education level) as a first step into the model did not significantly explain variance (R^2=.02, P=.09). Entering the symptom-related factors (stress level, depression, emotional exhaustion) and hope for improvement as a second block into the model resulted in significantly more explained variance (ΔR^2=.02, P=.046). Stress level and depression were predictive of treatment nonadherence in this step (t_{385}=2.00, P=.046; t_{385}=–2.10, P=.04). However, none of these factors was predictive of treatment nonadherence over and above the influence of guidance, which was added as dummy-coded variables (content-focused guidance vs administrative guidance; adherence-focused guidance vs administrative guidance) as a third step into the model. The final model explained 9.4% of the variance in treatment nonadherence (ΔR^2=.05, P<.001; t_{385}=4.26, P<.001; t_{385}=4.01, P<.001).

Table 2. Results of the imputed hierarchical multiple linear regression analysis to identify user characteristics predictive of treatment nonadherence over guidance formats.

Variable	R^2 or ΔR^{2a}	B (SE)[b]	Beta[c]	P
Step 1	.02			.09
Constant		3.22 (1.08)		
Gender		0.53 (0.31)	.09	.09
Education		0.47 (0.25)	.10	.06
Age		–0.00 (0.01)	–.01	.79
Step 2	.02			.046
Constant		0.28 (1.56)		.86
Gender		0.42 (0.31)	.07	.17
Education		0.44 (0.24)	.09	.07
Age		–0.01 (0.01)	–.03	.58
Stress		0.08 (0.04)	.14	.046
Depression		–0.04 (0.02)	–.15	.04
Emotional exhaustion		0.31 (0.21)	.09	.14
Hope for improvement		0.26 (0.20)	.07	.20
Step 3	.05			<.001
Constant		–0.49 (1.53)		
Gender		0.32 (0.30)	.05	.29
Education		0.44 (0.24)	.09	.06
Age		–0.01 (0.01)	–.04	.39
Stress		0.07 (0.04)	.12	.08
Depression		–0.03 (0.02)	–.11	.13
Emotional exhaustion		0.28 (0.20)	.08	.16
Hope for improvement		0.36 (0.20)	.09	.07
Guidance format (content-focused guidance vs administrative guidance)		1.31 (0.31)	.24	<.001
Guidance format (adherence-focused guidance vs administrative guidance)		1.24 (0.31)	.23	<.001

[a]R^2 is for step 1 and ΔR^2 for steps 2 and 3.

[b]Unstandardized regression coefficient and unstandardized standard error.

[c]Standardized regression coefficient.

XSL•FO

RenderX

Interactions Between Guidance Formats and Predictors

Adding the interactions between the guidance formats and the other variables to the model did not significantly change explained variance ($\Delta R^2=.02$, $P=.95$).

Discussion

Principal Results

The first aim of this study was to identify the adherence rates for an Internet-based mobile-supported stress management intervention. The content-focused guidance adherence rate was 71%, which is comparable to the rates found in other guided Internet-based stress management interventions (46%-88%) [21,22,83,84]. For the adherence-focused guidance and the administrative guidance conditions, we found adherence rates of 69% and 42%, respectively. Comparing adherence rates for adherence-focused guidance and administrative guidance to previous trials is difficult because most of those trials conducted in the field either used other guidance formats or did not report adherence rates. However, in a study with adherence reminders [85], a module completion rate of 35% was found, which is below the rate for adherence-focused guidance in this study (69%). Nevertheless, varying definitions and operationalization of module completion limit the comparability of treatment adherence rates between different studies. Despite the limited comparability, this study's results suggest relatively high adherence rates for content-focused guidance and adherence-focused guidance and lower rates for administrative guidance.

Our second research goal was to investigate the influence of different guidance formats on adherence in an Internet-based mobile-supported stress management intervention. Similarly to studies on other target conditions, such as depression [31] and anxiety [38,56], this study suggests that participants show better adherence with guided treatments in comparison to unguided treatments. As predicted, content-focused guidance and adherence-focused guidance resulted in higher treatment adherence rates compared to unguided treatment with only administrative support (administrative guidance).

As hypothesized, both content-focused guidance and adherence-focused guidance have high adherence rates; therefore, the next step was to analyze their equivalence in terms of treatment adherence. For both guidance formats, adherence was equivalent. Despite the equivalence in adherence rates, both guidance formats differ in the amount of eCoaching each requires. Content-focused guidance included both reminders and written feedback from an eCoach on every completed module and required up to 4 hours of coaching time. In contrast, adherence-focused guidance consisted of adherence monitoring and feedback on demand and only required up to 1 hour of coaching time per participant during the intervention. Therefore, by choosing adherence-focused guidance regarding the costs of treatment, substantial savings may be made without a significant reduction in patient adherence. This finding is in line with the assumption that the active factor responsible for improving adherence in guided versus unguided self-help interventions is that the participant is accompanied through the intervention. Providing instructions or detailed feedback on the content the participants worked on within the modules seems less critical for continued participant engagement.

However, the incremental value of offering feedback on demand compared to only adherence monitoring from an eCoach remains yet unclear. Within the adherence-focused guidance concept, it is hypothesized that feedback on demand is an important component so that the eCoach is seen as having the participant's best interests at heart. Offering support may be an antecedent for creating an adherence-promoting relationship and, according to the supportive accountability model [25], assumed to be necessary to maximize the effects of adherence monitoring. But, it is possible that the superiority of adherence-focused guidance compared to administrative guidance can be purely explained by the effect of personalized reminders. Thus, future studies should investigate whether having the possibility to contact an eCoach has an incremental influence on treatment adherence beyond adherence monitoring.

In the adherence-focused guidance study arm, monitoring adherence and sending a personalized reminder took up almost all the resources associated with this guidance format (approximately 1 hour per participant). In contrast, feedback on demand required much less resources. Hence, the question arises whether automatic reminders sent from the system on behalf of the eCoach have, in combination with feedback on demand, a similar effect on adherence, while requiring even less resources. Other studies have already shown positive effects on treatment adherence through automatic reminders [86].

If feedback on demand from a health professional is not a necessary component to achieve sufficient treatment adherence, adherence monitoring may also be performed by nonprofessionals. This would improve cost-effectiveness and dissemination. The eCoaches' qualification level was not found to significantly influence treatment efficacy in Internet interventions for a range of conditions [27]. A study by Titov and colleagues [87] did not show significant differences in clinical efficacy between layperson- or clinician-assisted Internet interventions. Hence, the characteristics of an eCoach keeping participants adherent to an intervention should be further investigated. It also remains unclear whether a personal coach is necessary at all.

However, reducing human contact in Internet interventions can also entail potential risks for participants. Without content-related feedback, the eCoach may not become aware of problems participants may experience during the training. Thus, the risk for negative effects with Internet interventions for some individuals has the potential to be higher when receiving adherence-focused guidance instead of content-focused guidance. For this reason, negative effects should be investigated in future studies that compare different guidance formats [88,89].

Likewise, the guidance format could also influence the acceptance and attractiveness of Internet interventions, and thereby be important for dissemination. Therefore, varying guidance formats in Internet interventions should also be

evaluated in terms of attractiveness and general acceptance [90,91].

The preceding discussion alludes to the many factors that can influence guidance in Internet interventions. Our third research question moves the focus from intervention characteristics to participant characteristics. It aimed to identify participant characteristics as predictors of nonadherence in an Internet-based mobile-supported stress management intervention. Although the guidance formats significantly predicted nonadherence in this study, the predictive value of the variables targeted in this analysis is small. The final model only explained 9.4% of the interindividual differences in nonadherence rates. Other variables, purported in previous literature as predictive of treatment nonadherence in Internet interventions, may prove to be incrementally important and should be investigated further. These include, for example, primary motivation, intention to adhere, and self-efficacy, as well as computer savviness, Internet affinity, and usability in addition to system- and program-related variables [25,44,92,93].

Limitations

This study has several limitations. First, the participants were not randomized to the three study arms. Hence, the differences between the guidance formats may be confounded by other differences between the studies.

Second, participants in the adherence-focused guidance group showed a significantly higher percentage of female users at baseline compared to participants in the administrative guidance group. Furthermore, the content-focused guidance sample had a significantly higher education level compared to the administrative guidance and adherence-focused guidance samples. These unsystematic variations may have contributed to better adherence rates in the adherence-focused guidance and content-focused guidance groups compared to the administrative guidance group. But gender and education level were not significantly associated with treatment nonadherence. However, the total sample was found to be highly educated, which may limit the generalization to the population.

Third, the adherence rates identified in this study refer to an Internet-based mobile-supported stress management intervention administered in RCTs. Several studies indicate that adherence rates to Internet interventions in the context of RCTs are higher than those available from open-access websites [7,9,94]. Improved usage rates in trial participants may indicate a preference for formal structures (eg, assessments, contact with research staff), which are thought to promote adherence. In addition, the elaborated study inclusion process (ie, completion of two self-report assessments, sending of informed consent) may have led to the inclusion of more motivated individuals through self-selection. As a result, adherence rates may not

apply to other Internet interventions or Internet interventions applied to a different context. Hence, treatment adherence rates in this study might not be generalizable to an Internet-based mobile-supported stress management intervention in an open-access practice. However, Internet interventions in open-access contexts lack the structure-giving elements of RCTs, such as contact with the study administration team. Hence, guidance could play an even more important role in open-access Internet interventions by structuring the course of treatment. Thus, the differences between content-focused guidance, adherence-focused guidance, and administrative guidance may have been underestimated in this study.

Fourth, generalizability of the treatment adherence rates in an Internet-based mobile-supported stress management intervention are further limited by only including employees with elevated stress levels in the studies. Thus, the findings of this study may not be relevant for settings with participants that were not preselected based on their stress level. However, the stress level was not significantly associated with lower adherence rates beyond the guidance formats.

Fifth, only one adherence measure was included in the analysis. Different adherence measures need to be used which capture the quality of engagement with an intervention to a greater extent, such as time on website, number of completed homework assignments, and diary entries [95].

Sixth, as in most predictor studies, the analyses in this study were exploratory without any presumptions about the relationship between the predictors and adherence in order to generate hypotheses [96].

Conclusion and Future Research

This study has important implications for research and practice. Guidance with focus on treatment adherence has the potential to be helpful in keeping participants involved in the training and, at the same time, keeping coaching costs low. Evaluating the cost-effectiveness as well as the comparative efficacy of the different guidance formats should be a next step to further complement the findings of this study [40]. Moreover, low treatment adherence is not negative per se because participants might quit the intervention due to improvement or because it does not meet their needs. Therefore, it is important to identify the reasons for nonadherence qualitatively to assess the impact of nonadherence for the efficacy of an intervention. Future research should focus on developing a theoretical model to explain the mechanisms behind treatment nonadherence in Internet interventions and derive variables from existing theories. This information can be used as a framework for the construction of an algorithm that generates risk profiles for participants that may be more susceptible to nonadherence.

Conflicts of Interest

Dirk Lehr, Matthias Berking and David Ebert are stakeholders of the "Institute for Online Health Trainings", that aims to transfer scientific knowledge related to the present research into routine healthcare.

References

1. Batterham PJ, Neil AL, Bennett K, Griffiths KM, Christensen H. Predictors of adherence among community users of a cognitive behavior therapy website. Patient Prefer Adherence 2008;2:97-105 [FREE Full text] [Medline: 19920949]

2. Bossen D, Buskermolen M, Veenhof C, de Bakker BD, Dekker J. Adherence to a web-based physical activity intervention for patients with knee and/or hip osteoarthritis: a mixed method study. J Med Internet Res 2013;15(10):e223 [FREE Full text] [doi: 10.2196/jmir.2742] [Medline: 24132044]

3. Calear AL, Christensen H, Mackinnon A, Griffiths KM. Adherence to the MoodGYM program: outcomes and predictors for an adolescent school-based population. J Affect Disord 2013 May;147(1-3):338-344. [doi: 10.1016/j.jad.2012.11.036] [Medline: 23245469]

4. Farvolden P, Denisoff E, Selby P, Bagby RM, Rudy L. Usage and longitudinal effectiveness of a Web-based self-help cognitive behavioral therapy program for panic disorder. J Med Internet Res 2005;7(1):e7 [FREE Full text] [doi: 10.2196/jmir.7.1.e7] [Medline: 15829479]

5. Newby JM, Mackenzie A, Williams AD, McIntyre K, Watts S, Wong N, et al. Internet cognitive behavioural therapy for mixed anxiety and depression: a randomized controlled trial and evidence of effectiveness in primary care. Psychol Med 2013 Dec;43(12):2635-2648. [doi: 10.1017/S0033291713000111] [Medline: 23419552]

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. 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]

8. Beintner I, Jacobi C, Taylor CB. Participant adherence to the Internet-based prevention program StudentBodies™ for eating disorders — A review. Internet Interventions 2014 Mar;1(1):26-32. [doi: 10.1016/j.invent.2014.03.001]

9. Christensen H, Griffiths KM, Korten AE, Brittliffe K, Groves C. A comparison of changes in anxiety and depression symptoms of spontaneous users and trial participants of a cognitive behavior therapy website. J Med Internet Res 2004 Dec 22;6(4):e46 [FREE Full text] [doi: 10.2196/jmir.6.4.e46] [Medline: 15631970]

10. 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]

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

12. Manwaring JL, Bryson SW, Goldschmidt AB, Winzelberg AJ, Luce KH, Cunning D, et al. Do adherence variables predict outcome in an online program for the prevention of eating disorders? J Consult Clin Psychol 2008 Apr;76(2):341-346. [doi: 10.1037/0022-006X.76.2.341] [Medline: 18377129]

13. Meyer B, Berger T, Caspar F, Beevers CG, Andersson G, Weiss M. Effectiveness of a novel integrative online treatment for depression (Deprexis): randomized controlled trial. J Med Internet Res 2009;11(2):e15 [FREE Full text] [doi: 10.2196/jmir.1151] [Medline: 19632969]

14. Stansfeld S, Candy B. Psychosocial work environment and mental health--a meta-analytic review. Scand J Work Environ Health 2006 Dec;32(6):443-462 [FREE Full text] [Medline: 17173201]

15. Billings DW, Cook RF, Hendrickson A, Dove DC. A web-based approach to managing stress and mood disorders in the workforce. J Occup Environ Med 2008 Aug;50(8):960-968. [doi: 10.1097/JOM.0b013e31816c435b] [Medline: 18695455]

16. Ebert DD, Lehr D, Boß L, Riper H, Cuijpers P, Andersson G, et al. Efficacy of an internet-based problem-solving training for teachers: results of a randomized controlled trial. Scand J Work Environ Health 2014 Nov;40(6):582-596. [doi: 10.5271/sjweh.3449] [Medline: 25121986]

17. Eisen KP, Allen GJ, Bollash M, Pescatello LS. Stress management in the workplace: a comparison of a computer-based and an in-person stress-management intervention. Comput Hum Behav 2008 Mar;24(2):486-496. [doi: 10.1016/j.chb.2007.02.003]

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

19. Hasson D, Anderberg UM, Theorell T, Arnetz BB. Psychophysiological effects of a web-based stress management system: a prospective, randomized controlled intervention study of IT and media workers [ISRCTN54254861]. BMC Public Health 2005;5:78 [FREE Full text] [doi: 10.1186/1471-2458-5-78] [Medline: 16042796]

20.   Drozd F, Raeder S, Kraft P, Bjørkli CA. Multilevel growth curve analyses of treatment effects of a Web-based intervention
       for stress reduction: randomized controlled trial. J Med Internet Res 2013;15(4):e84 [FREE Full text] [doi: 10.2196/jmir.2570]
       [Medline: 23607962]

21.   Hänggi Y. Kann web-basierte Stressprävention erfolgreich sein? Zeitschrift für Klinische Psychologie und Psychotherapie
       2006 Jul 1;35(3):169-177. [doi: 10.1026/1616-3443.35.3.169]

22.   Zetterqvist K, Maanmies J, Ström L, Andersson G. Randomized controlled trial of internet-based stress management. Cogn
       Behav Ther 2003;32(3):151-160. [doi: 10.1080/16506070302316] [Medline: 16291546]

23.   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]

24.   Beintner I, Jacobi C, Schmidt UH. Participation and outcome in manualized self-help for bulimia nervosa and binge eating
       disorder - a systematic review and metaregression analysis. Clin Psychol Rev 2014 Mar;34(2):158-176. [doi:
       10.1016/j.cpr.2014.01.003] [Medline: 24508686]

25.   Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence
       to eHealth interventions. J Med Internet Res 2011;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]

26.   Mohr DC, Duffecy J, Ho J, Kwasny M, Cai X, Burns MN, et al. A randomized controlled trial evaluating a manualized
       TeleCoaching protocol for improving adherence to a web-based intervention for the treatment of depression. PLoS One
       2013 Aug;8(8):e70086 [FREE Full text] [doi: 10.1371/journal.pone.0070086] [Medline: 23990896]

27.   Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on Internet-based mental health interventions —
       a systematic review. Internet Interventions 2014 Oct;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003]

28.   Berger T, Hämmerli K, Gubser N, Andersson G, Caspar F. Internet-based treatment of depression: a randomized controlled
       trial comparing guided with unguided self-help. Cogn Behav Ther 2011 Dec;40(4):251-266. [doi:
       10.1080/16506073.2011.616531] [Medline: 22060248]

29.   Nicholas J, Proudfoot J, Parker G, Gillis I, Burckhardt R, Manicavasagar V, et al. The ins and outs of an online bipolar
       education program: a study of program attrition. J Med Internet Res 2010;12(5):e57 [FREE Full text] [doi: 10.2196/jmir.1450]
       [Medline: 21169169]

30.   Proudfoot J, Parker G, Manicavasagar V, Hadzi-Pavlovic D, Whitton A, Nicholas J, et al. Effects of adjunctive peer support
       on perceptions of illness control and understanding in an online psychoeducation program for bipolar disorder: a randomised
       controlled trial. J Affect Disord 2012 Dec 15;142(1-3):98-105. [doi: 10.1016/j.jad.2012.04.007] [Medline: 22858215]

31.   Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis.
       Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]

32.   van Ballegooijen W, Cuijpers P, van Straten A, Karyotaki E, Andersson G, Smit JH, et al. Adherence to Internet-based and
       face-to-face cognitive behavioural therapy for depression: a meta-analysis. PLoS One 2014;9(7):e100674 [FREE Full text]
       [doi: 10.1371/journal.pone.0100674] [Medline: 25029507]

33.   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]

34.   Newman MG, Szkodny LE, Llera SJ, Przeworski A. A review of technology-assisted self-help and minimal contact therapies
       for anxiety and depression: is human contact necessary for therapeutic efficacy? Clin Psychol Rev 2011 Feb;31(1):89-103.
       [doi: 10.1016/j.cpr.2010.09.008] [Medline: 21130939]

35.   Clarke G, Reid E, Eubanks D, O'Connor E, DeBar LL, Kelleher C, et al. Overcoming depression on the Internet (ODIN):
       a randomized controlled trial of an Internet depression skills intervention program. J Med Internet Res 2002 Dec;4(3):E14
       [FREE Full text] [doi: 10.2196/jmir.4.3.e14] [Medline: 12554545]

36.   Clarke G, Eubanks D, Reid E, Kelleher C, O'Connor E, DeBar LL, et al. Overcoming Depression on the Internet (ODIN)
       (2): a randomized trial of a self-help depression skills program with reminders. J Med Internet Res 2005;7(2):e16 [FREE
       Full text] [doi: 10.2196/jmir.7.2.e16] [Medline: 15998607]

37.   Schneider F, van Osch L, Schulz DN, Kremers SP, de Vries H. The influence of user characteristics and a periodic email
       prompt on exposure to an internet-delivered computer-tailored lifestyle program. J Med Internet Res 2012;14(2):e40 [FREE
       Full text] [doi: 10.2196/jmir.1939] [Medline: 22382037]

38.   Titov N, Andrews G, Choi I, Schwencke G, Johnston L. Randomized controlled trial of web-based treatment of social
       phobia without clinician guidance. Aust NZ J Psychiatry 2009 Jan;43(10):913-919. [doi: 10.1080/00048670903179160]

39.   Titov N, Dear BF, Johnston L, Lorian C, Zou J, Wootton B, et al. Improving adherence and clinical outcomes in self-guided
       internet treatment for anxiety and depression: randomised controlled trial. PLoS One 2013;8(7):e62873 [FREE Full text]
       [doi: 10.1371/journal.pone.0062873] [Medline: 23843932]

40.   Ebert DD, Lehr D, Smit F, Zarski A, Riper H, Heber E, et al. Efficacy and cost-effectiveness of minimal guided and unguided
       internet-based mobile supported stress-management in employees with occupational stress: a three-armed randomised
       controlled trial. BMC Public Health 2014;14:807 [FREE Full text] [doi: 10.1186/1471-2458-14-807] [Medline: 25099533]

41.   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 [FREE Full text] [doi: 10.1007/s12160-009-9133-4] [Medline:
       19802647]

42.   Murray E, White IR, Varagunam M, Godfrey C, Khadjesari Z, McCambridge J. Attrition revisited: adherence and retention
      in a web-based alcohol trial. J Med Internet Res 2013;15(8):e162 [FREE Full text] [doi: 10.2196/jmir.2336] [Medline:
      23996958]

43.   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]

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

45.   Zbikowski SM, Hapgood J, Smucker BS, McAfee T. Phone and web-based tobacco cessation treatment: real-world utilization
      patterns and outcomes for 11,000 tobacco users. J Med Internet Res 2008;10(5):e41 [FREE Full text] [doi: 10.2196/jmir.999]
      [Medline: 19017583]

46.   Lange A, Rietdijk D, Hudcovicova M, van de Ven JP, Schrieken B, Emmelkamp PM. Interapy: a controlled randomized
      trial of the standardized treatment of posttraumatic stress through the internet. J Consult Clin Psychol 2003 Oct;71(5):901-909.
      [doi: 10.1037/0022-006X.71.5.901] [Medline: 14516238]

47.   Linke S, Murray E, Butler C, Wallace P. Internet-based interactive health intervention for the promotion of sensible drinking:
      patterns of use and potential impact on members of the general public. J Med Internet Res 2007;9(2):e10 [FREE Full text]
      [doi: 10.2196/jmir.9.2.e10] [Medline: 17513281]

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

49.   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]

50.   Spek V, Nyklíček I, Smits N, Cuijpers P, Riper H, Keyzer J, et al. Internet-based cognitive behavioural therapy for
      subthreshold depression in people over 50 years old: a randomized controlled clinical trial. Psychol Med 2007
      Dec;37(12):1797-1806. [doi: 10.1017/S0033291707000542] [Medline: 17466110]

51.   Spek V, Nyklíček I, Cuijpers P, Pop V. Predictors of outcome of group and Internet-based cognitive behavior therapy. J
      Affect Disord 2008 Jan;105(1-3):137-145. [doi: 10.1016/j.jad.2007.05.001] [Medline: 17543392]

52.   Neil AL, Batterham P, Christensen H, Bennett K, Griffiths KM. Predictors of adherence by adolescents to a cognitive
      behavior therapy website in school and community-based settings. J Med Internet Res 2009;11(1):e6 [FREE Full text] [doi:
      10.2196/jmir.1050] [Medline: 19275982]

53.   Postel MG, de Haan HA, ter Huurne ED, van der Palen J, Becker ES, de Jong CA. Attrition in web-based treatment for
      problem drinkers. J Med Internet Res 2011;13(4):e117 [FREE Full text] [doi: 10.2196/jmir.1811] [Medline: 22201703]

54.   Farrer LM, Griffiths KM, Christensen H, Mackinnon AJ, Batterham PJ. Predictors of adherence and outcome in Internet-based
      cognitive behavior therapy delivered in a telephone counseling Setting. Cogn Ther Res 2013 Oct 27;38(3):358-367. [doi:
      10.1007/s10608-013-9589-1]

55.   von Brachel R, Hötzel K, Hirschfeld G, Rieger E, Schmidt U, Kosfelder J, et al. Internet-based motivation program for
      women with eating disorders: eating disorder pathology and depressive mood predict dropout. J Med Internet Res
      2014;16(3):e92 [FREE Full text] [doi: 10.2196/jmir.3104] [Medline: 24686856]

56.   Nordgreen T, Havik OE, Ost LG, Furmark T, Carlbring P, Andersson G. Outcome predictors in guided and unguided
      self-help for social anxiety disorder. Behav Res Ther 2012 Jan;50(1):13-21. [doi: 10.1016/j.brat.2011.10.009] [Medline:
      22134140]

57.   Wojtowicz M, Day V, McGrath PJ. Predictors of participant retention in a guided online self-help program for university
      students: prospective cohort study. J Med Internet Res 2013;15(5):e96 [FREE Full text] [doi: 10.2196/jmir.2323] [Medline:
      23697614]

58.   Boettcher J, Renneberg B, Berger T. Patient expectations in internet-based self-help for social anxiety. Cogn Behav Ther
      2013;42(3):203-214. [doi: 10.1080/16506073.2012.759615] [Medline: 23697570]

59.   Cavanagh K, Shapiro DA, Van Den Berg S, Swain S, Barkham M, Proudfoot J. The acceptability of computer-aided
      cognitive behavioural therapy: a pragmatic study. Cogn Behav Ther 2009;38(4):235-246. [doi: 10.1080/16506070802561256]
      [Medline: 19306147]

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

61.   Heber E, Ebert DD, Lehr D, Nobis S, Berking M, Riper H. Efficacy and cost-effectiveness of a web-based and mobile
      stress-management intervention for employees: design of a randomized controlled trial. BMC Public Health 2013;13:655
      [FREE Full text] [doi: 10.1186/1471-2458-13-655] [Medline: 23855376]

62.   Heber E, Lehr D, Ebert DD, Berking M, Riper H. Web-based and mobile stress management intervention for employees:
      a randomized controlled trial. J Med Internet Res 2016;18(1):e21 [FREE Full text] [doi: 10.2196/jmir.5112] [Medline:
      26818683]

63. Ebert DD, Heber E, Berking M, Riper H, Cuijpers P, Funk B, et al. Self-guided internet-based and mobile-based stress management for employees: results of a randomised controlled trial. Occup Environ Med 2016 May;73(5):315-323. [doi: 10.1136/oemed-2015-103269] [Medline: 26884049]

64. Ebert DD, Heber E, Riper H, Cuijpers P, Berking M, Lehr D. Efficacy and mechanism of change of Internet- and mobile based stress management for employees with adherence focused guidance: a randomized controlled trial. Scand Journ Work Env Health 2016:tbd (forthcoming).

65. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]

66. Beck AT, Steer AS, Brown GK. BDI-II: Beck Depression Inventory Manual. 2nd ed. San Antonio, TX: Psychological Corporation; 1996.

67. Junge MN, Lehr D, Bockting CL, Berking M, Riper H, Cuijpers P, et al. For whom are internet-based occupational mental health interventions effective? Moderators of internet-based problem-solving training outcome. Internet Interventions 2015 Mar;2(1):39-47. [doi: 10.1016/j.invent.2014.11.007]

68. Nezu AM, Perri MG. Social problem-solving therapy for unipolar depression: an initial dismantling investigation. J Consult Clin Psychol 1989 Jun;57(3):408-413. [Medline: 2738213]

69. 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]

70. Berking M, Whitley B. Affect Regulation Training. New York: Springer New York; 2014.

71. Berking M, Ebert D, Cuijpers P, Hofmann SG. Emotion regulation skills training enhances the efficacy of inpatient cognitive behavioral therapy for major depressive disorder: a randomized controlled trial. Psychother Psychosom 2013;82(4):234-245. [doi: 10.1159/000348448] [Medline: 23712210]

72. Michalsen A, Jeitler M, Brunnhuber S, Lüdtke R, Büssing A, Musial F, et al. Iyengar yoga for distressed women: a 3-armed randomized controlled trial. Evid Based Complement Alternat Med 2012;2012:408727 [FREE Full text] [doi: 10.1155/2012/408727] [Medline: 23049608]

73. Hautzinger M, Bailer M, Hofmeister D, Keller F. Allgemeine Depressions Skala (ADS) Manual [Center for Epidemiological Studies Depression Scale (CES-D; Radloff LS, 1977)]. Boston, MA: Hogrefe Publishers; 2012.

74. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Appl Psych Meas 1977 Jun 01;1(3):385-401. [doi: 10.1177/014662167700100306]

75. Neubach B, Schmidt K. Gütekriterien einer deutschen Fassung des Mastach Burnout Inventory (MBI-D) - Eine Replikationsstudie bei Altenpflegekräften. Zeitschrift für Arbeits- und Organisationspsychologie A&O 2000 Jul 1;44(3):140-156. [doi: 10.1026//0932-4089.44.3.140]

76. Schaufeli WB, Leiter MP, Maslach C, Jackson SE. The MBI-general survey. In: Maslach C, Jackson SE, Leiter MP, editors. Maslach Burnout Inventory Manual. 3rd ed. Palo Alto, CA: Consulting Psychologists Press; 1996:19-26.

77. Schulte D. Messung der Therapieerwartung und Therapieevaluationvon Patienten (PATHEV). Zeitschrift für Klinische Psychologie und Psychotherapie 2005 Jul 1;34(3):176-187. [doi: 10.1026/1616-3443.34.3.176]

78. Mertler CA, Vannatta RA. Advanced and Multivariate Statistical Methods: Practical Application and Interpretation (5th ed). Glendale, CA: Pyrczak Publishing; 2000.

79. Kirk RE. Practical significance: a concept whose time has come. Educ Psychol Meas 1996 Oct 01;56(5):746-759. [doi: 10.1177/0013164496056005002]

80. Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [Medline: 19565683]

81. Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. JAMA 2012 Dec 26;308(24):2594-2604. [doi: 10.1001/jama.2012.87802] [Medline: 23268518]

82. Wiens BL. Choosing an equivalence limit for noninferiority or equivalence studies. Control Clin Trials 2002 Feb;23(1):2-14. [Medline: 11852160]

83. Rose RD, Buckey JC, Zbozinek TD, Motivala SJ, Glenn DE, Cartreine JA, et al. A randomized controlled trial of a self-guided, multimedia, stress management and resilience training program. Behav Res Ther 2013 Feb;51(2):106-112. [doi: 10.1016/j.brat.2012.11.003] [Medline: 23262118]

84. Ruwaard J, Lange A, Bouwman M, Broeksteeg J, Schrieken B. E-mailed standardized cognitive behavioural treatment of work-related stress: a randomized controlled trial. Cogn Behav Ther 2007;36(3):179-192. [doi: 10.1080/16506070701381863] [Medline: 17852171]

85. Morledge TJ, Allexandre D, Fox E, Fu AZ, Higashi MK, Kruzikas DT, et al. Feasibility of an online mindfulness program for stress management: a randomized, controlled trial. Ann Behav Med 2013 Oct;46(2):137-148 [FREE Full text] [doi: 10.1007/s12160-013-9490-x] [Medline: 23632913]

86. Robertson L, Smith M, Castle D, Tannenbaum D. Using the Internet to enhance the treatment of depression. R Aust New Zeal Coll Psychiatr 2006;14(4):413-417. [Medline: 18921256]

87. Titov N, Andrews G, Davies M, McIntyre K, Robinson E, Solley K. Internet treatment for depression: a randomized controlled trial comparing clinician vs. technician assistance. PLoS One 2010;5(6):e10939 [FREE Full text] [doi: 10.1371/journal.pone.0010939] [Medline: 20544030]

88. Rozental A, Andersson G, Boettcher J, Ebert DD, Cuijpers P, Knaevelsrud C, et al. Consensus statement on defining and measuring negative effects of Internet interventions. Internet Interventions 2014 Mar;1(1):12-19. [doi: 10.1016/j.invent.2014.02.001]

89. Ebert DD, Lehr D, Baumeister H, Boß L, Riper H, Cuijpers P, et al. GET.ON Mood Enhancer: efficacy of Internet-based guided self-help compared to psychoeducation for depression: an investigator-blinded randomised controlled trial. Trials 2014;15:39 [FREE Full text] [doi: 10.1186/1745-6215-15-39] [Medline: 24476555]

90. Baumeister H, Nowoczin L, Lin J, Seifferth H, Seufert J, Laubner K, et al. Impact of an acceptance facilitating intervention on diabetes patients' acceptance of Internet-based interventions for depression: a randomized controlled trial. Diabetes Res Clin Pract 2014 Jul;105(1):30-39. [doi: 10.1016/j.diabres.2014.04.031] [Medline: 24862240]

91. Baumeister H, Seifferth H, Lin J, Nowoczin L, Lüking M, Ebert D. Impact of an acceptance facilitating intervention on patients' acceptance of Internet-based pain interventions: a randomized controlled trial. Clin J Pain 2015 Jun;31(6):528-535. [doi: 10.1097/AJP.0000000000000118] [Medline: 24866854]

92. Hebert EA, Vincent N, Lewycky S, Walsh K. Attrition and adherence in the online treatment of chronic insomnia. Behav Sleep Med 2010;8(3):141-150. [doi: 10.1080/15402002.2010.487457] [Medline: 20582757]

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

94. Poston JM, Hanson WE. Meta-analysis of psychological assessment as a therapeutic intervention. Psychol Assess 2010 Jun;22(2):203-212. [doi: 10.1037/a0018679] [Medline: 20528048]

95. Mohr DC, Burns MN, Schueller SM, Clarke G, Klinkman M. Behavioral intervention technologies: evidence review and recommendations for future research in mental health. Gen Hosp Psychiatry 2013 Aug;35(4):332-338 [FREE Full text] [doi: 10.1016/j.genhosppsych.2013.03.008] [Medline: 23664503]

96. Kraemer HC, Wilson GT, Fairburn CG, Agras WS. Mediators and moderators of treatment effects in randomized clinical trials. Arch Gen Psychiatry 2002;59(10):877-883. [Medline: 12365874]

Abbreviations

RCT: randomized controlled trial

Original Paper

A Dietary Feedback System for the Delivery of Consistent Personalized Dietary Advice in the Web-Based Multicenter Food4Me Study

Hannah Forster[1], BSc (Hons); Marianne C Walsh[1], PhD; Clare B O'Donovan[1], MSc; Clara Woolhead[1], MSc; Caroline McGirr[1], BSc (Hons); E.J Daly[2], MSc; Richard O'Riordan[2], PGDip; Carlos Celis-Morales[3], PhD; Rosalind Fallaize[4], PhD; Anna L Macready[4], PhD; Cyril F M Marsaux[5], MSc; Santiago Navas-Carretero[6], PhD; Rodrigo San-Cristobal[6], MSc; Silvia Kolossa[7], MSc; Kai Hartwig[7], PhD; Christina Mavrogianni[8], MSc; Lydia Tsirigoti[8], MSc; Christina P Lambrinou[8], MSc; Magdalena Godlewska[9], MSc; Agnieszka Surwiłło[9], MSc; Ingrid Merethe Fange Gjelstad[10], PhD; Christian A Drevon[10], PhD; Yannis Manios[8], PhD; Iwona Traczyk[9], PhD; J Alfredo Martinez[6], PhD; Wim H M Saris[5], PhD; Hannelore Daniel[7], PhD; Julie A Lovegrove[4], PhD; John C Mathers[3], PhD; Michael J Gibney[1], PhD; Eileen R Gibney[1], PhD; Lorraine Brennan[1], PhD

[1]UCD Institute of Food and Health, University College Dublin, Dublin, Ireland

[2]Creme Global, Trinity Technology and Enterprise Campus, Grand Canal Quay, Dublin, Ireland

[3]Human Nutrition Research Centre, Institute of Cellular Medicine, Newcastle University, Newcastle Upon Tyne, United Kingdom

[4]Hugh Sinclair Unit of Human Nutrition and Institute for Cardiovascular and Metabolic Research, University of Reading, Reading, United Kingdom

[5]Department of Human Biology, NUTRIM, School for Nutrition and Translational Research in Metabolism, Maastricht University Medical Centre, Maastricht, Netherlands

[6]Department of Nutrition, Food Science and Physiology, Centre for Nutrition Research, University of Navarra, Pamplona, Spain

[7]ZIEL Research Center of Nutrition and Food Sciences, Biochemistry Unit, Technische Universität München, München, Germany

[8]Department of Nutrition and Dietetics, Harokopio University, Athens, Greece

[9]National Food & Nutrition Institute (IZZ), Warsaw, Poland

[10]Department of Nutrition, Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway

Corresponding Author:
Lorraine Brennan, PhD
UCD Institute of Food and Health
University College Dublin
Belfield
Dublin,
Ireland
Phone: 353 1 716 ext 2811
Fax: 353 1 716147
Email: lorraine.brennan@ucd.ie

Abstract

Background: Despite numerous healthy eating campaigns, the prevalence of diets high in saturated fatty acids, sugar, and salt and low in fiber, fruit, and vegetables remains high. With more people than ever accessing the Internet, Web-based dietary assessment instruments have the potential to promote healthier dietary behaviors via personalized dietary advice.

Objective: The objectives of this study were to develop a dietary feedback system for the delivery of consistent personalized dietary advice in a multicenter study and to examine the impact of automating the advice system.

Methods: The development of the dietary feedback system included 4 components: (1) designing a system for categorizing nutritional intakes; (2) creating a method for prioritizing 3 nutrient-related goals for subsequent targeted dietary advice; (3) constructing decision tree algorithms linking data on nutritional intake to feedback messages; and (4) developing personal feedback reports. The system was used manually by researchers to provide personalized nutrition advice based on dietary assessment to 369 participants during the Food4Me randomized controlled trial, with an automated version developed on completion of the study.

XSL•FO
RenderX

Results: Saturated fatty acid, salt, and dietary fiber were most frequently selected as nutrient-related goals across the 7 centers. Average agreement between the manual and automated systems, in selecting 3 nutrient-related goals for personalized dietary advice across the centers, was highest for nutrient-related goals 1 and 2 and lower for goal 3, averaging at 92%, 87%, and 63%, respectively. Complete agreement between the 2 systems for feedback advice message selection averaged at 87% across the centers.

Conclusions: The dietary feedback system was used to deliver personalized dietary advice within a multi-country study. Overall, there was good agreement between the manual and automated feedback systems, giving promise to the use of automated systems for personalizing dietary advice.

Trial Registration: Clinicaltrials.gov NCT01530139; https://clinicaltrials.gov/ct2/show/NCT01530139 (Archived by WebCite at http://www.webcitation.org/6ht5Dgj8I)

KEYWORDS

Introduction

Diets that are low in saturated fatty acids (SFAs), sugar, and salt and high in fruit, vegetables, and fiber are considered the healthy choice and have been shown to reduce the risk of noncommunicable diseases (NCD) [1-4]. However, despite numerous campaigns and policies to promote healthy eating, NCD burden has continued to rise, globally, over the past decade with increased contribution from nutrition-related risk factors [5]. As a result, there is the need for effective strategies to promote healthy dietary habits and to help consumers to achieve the necessary dietary changes.

Intensive inter-person counseling and interventions have been shown to improve dietary behaviors [6], although the potential feasibility and effectiveness of such methods across large populations is limited by both expense and accessibility [1,7]. Given the global increases in Internet availability, and the increasing utilization of the Internet as a method for delivering behavioral changes [8-11], the use of Web-based dietary assessment tools to provide personalized nutrition (PN) advice (based on dietary intake alone) at an individual level could provide a more cost-effective approach for improving dietary behaviors than current generic dietary advice and inter-person counseling [1,12-14].

Numerous studies have shown that personalized feedback advice is more effective than generic information for changing health behaviors including dietary intake [8,15-20]. Web-based dietary assessment instruments (eg, food frequency questionnaires—FFQs, 24-hour recalls, food diaries) can be the basis for the development of tools with the ability to provide users with personalized or tailored dietary feedback advice based on their self-reported dietary habits [12]. Such dietary feedback advice often illustrates the adequacy of the individual's nutrient intakes in comparison with recommended intakes, intakes of peers of the same age or gender or previously attained intakes [7,16,19-21].

Although Web-based dietary assessment tools provide a potential starting point for delivering personalized dietary advice to large populations, they are usually 1-way systems that are designed to collect dietary intake data in a cost-effective manner. To limit measurement error and increase data completeness,

Web-based dietary assessment tools can be preprogrammed with range and plausibility checks, probing questions and encoded to ensure all questions are answered [22]. Despite programming these check elements, it is essential that dietary assessment systems are coherent, comprehensive, structured, and clear, particularly when a consistent approach for composing and delivering dietary feedback advice is required. Algorithms can be developed to provide feedback in a systematic manner and have been applied in many clinical or health care domains to link personal data to knowledge systems with the aim of developing diagnoses or treatment options. In the same way, by comparing nutrient intakes with dietary recommendations, it is possible to develop algorithms that generate appropriate feedback messages, which can be stored in a message archive [23,24]. Algorithms can also be programmed into dietary assessment tools, facilitating the automatic generation of feedback advice [8]. Web-based dietary assessment tools, which generate automatic feedback advice, have the potential to improve dietary habits across large population groups, while minimizing researcher burden, reducing costs, and saving time.

The objectives of this paper were to report on the development of a Web-based dietary feedback system for the delivery of consistent personalized dietary advice in a multicenter study and to examine the impact of automating the feedback system on dietary advice delivery.

Methods

Development of the Feedback System

The dietary feedback system was developed, for manual use by researchers, for the delivery of personalized dietary advice in the pan-European Food4Me Proof-of-Principle (PoP) study. The Food4Me PoP study that aimed to deliver Web-based personalized dietary and physical activity (PA) advice was designed to emulate a real-life Internet-delivered, PN service [25]. As described in detail elsewhere, 1607 participants were randomized to 1 of the 4 groups receiving different levels of PN advice: Level 0 (control group) receiving conventional, non-PN advice; Level 1 PN advice based on dietary intake and PA data alone; Level 2 PN advice based on dietary intake, PA, and phenotypic data; Level 3 PN advice based on dietary intake, PA, phenotypic, and genotypic data [25]. The aim was to recruit

a total of 1540 participants (220 participants per center) to allow for a potential 20% dropout rate, planned using a priori power calculation [25]. Power calculations were conducted in Minitab and based on glucose and omega-3 fatty acid concentrations within European adult populations [25]. The study had ethical approval from the corresponding committees of all participating centers (Germany, Greece, Ireland, the Netherlands, Poland, Spain, and the United Kingdom) and was conducted from August 2012 to March 2014. The Food4Me trial was registered as a randomized control trial (NCT01530139) at Clinicaltrials.gov.

The dietary feedback system, described in this paper, was used to deliver PN advice to participants randomized to receive Level 1 PN only (n=414). All participants of the Food4Me study received dietary feedback advice via email without face-to-face contact with researchers [25]. Thus, the feedback system was designed and developed to ensure that delivery of personalized dietary feedback was consistent across all the 7 countries. The key stages in the development of the dietary feedback system are illustrated in Figure 1 and described in detail in the following section.

Figure 1. Stages in the development of the dietary feedback system.
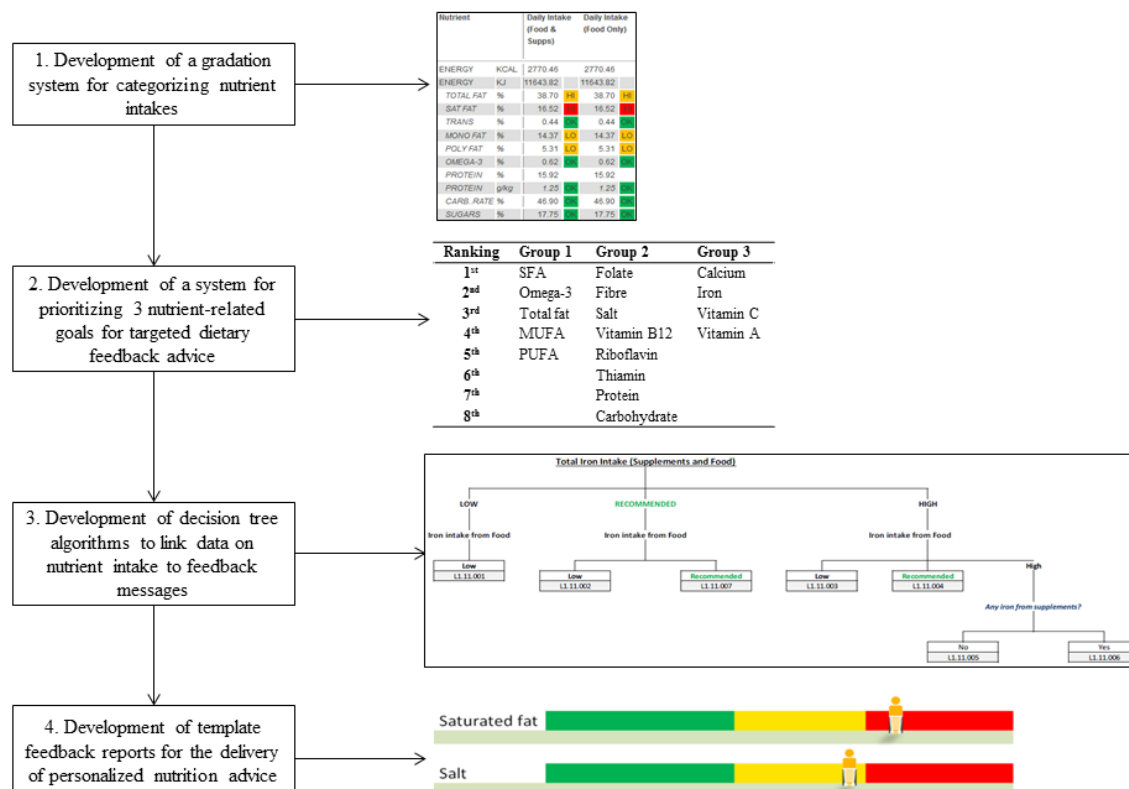
Development of a Gradation System for Coding of Nutrient Intakes

Dietary intake data were collected throughout the Food4Me study using the recently developed and validated Web-based Food4Me FFQ [26,27]. Previous evidence has shown the Food4Me FFQ to have good agreement with the EPIC-Norfolk FFQ and moderate agreement with a 4-day weighed food record for the assessment of both nutrient and food group intakes, rendering it a useful tool for ranking individuals based on nutrient and food group intakes [26,27]. The Institute of Medicine (IOM) dietary reference intakes were used as a basis for developing a gradation system to categorize nutrient intakes automatically after completion of the Web-based Food4Me FFQ by participants [28,29]. Institute of Medicine reference values were used because (1) they were the most up-to-date values available and (2) recommendations vary across Europe with a need to find standardized reference intakes. Nutrient gradations

were calculated based on IOM estimated average requirements (EARs) and tolerable upper intake levels (upper limits—ULs). Lower cutoff values were calculated as the EAR minus 2 standard deviations. Institute of Medicine–recommended daily allowance or World Health Organization recommendations were used to calculate the gradations when IOM EARs were not available [30,31]. The gradations (EAR, UL, and lower cutoff point) for each nutrient were preprogrammed into the Web-based Food4Me FFQ by Creme Global (Dublin, Ireland) facilitating a visual ranking system for nutrient intakes (Multimedia Appendix 1). Where applicable, gradations incorporated IOM recommendations for age and gender. Color coding and labeling of nutrient intakes were integrated into the Web-based Food4Me FFQ nutritional outputs to enable rapid visual assessment of the nutrient intake as *"very low," "low," "recommended," "high,"* or *"very high,"* which were labeled and color-coded as red, amber, green, amber, and red, respectively.

Development of a System for Prioritizing Dietary Feedback Advice

Seventeen nutrients were selected for inclusion in the dietary feedback system viz *"protein," "carbohydrate," "total fat," "MUFA," "PUFA," "SFA," "salt," "omega 3," "fiber," "calcium," "iron," "vitamin A," "folate," "thiamin," "riboflavin," "vitamin B12,"* and *"vitamin C"*. In addition, based on patient-centered models for facilitating dietary change, which emphasize that patients should focus on only a few goals at a time [25], 3 nutrient-related goals (target nutrients) were selected for particular emphasis in the feedback report. To select these 3 nutrient-related goals for subsequent dietary advice, the 17 nutrients were split into 3 groups (Multimedia Appendix 2). Group 1 consisted of all the "fat-related" nutrients: SFA, omega-3, total fat, monounsaturated fatty acid, and polyunsaturated fatty acid (PUFA). Group 2 included folate, dietary fiber, salt, vitamin B12, riboflavin, thiamin, protein, and carbohydrate. Group 3 consisted of calcium, iron, vitamin C, and vitamin A. A ranking system was embedded in the methodology for identifying target nutrients, so that nutrients at the top of each group received highest priority (ie, nutrients of higher public health concern). Generally, the highest priority nutrient, flagged "red" from each group was chosen as the nutrient-related goal, if no "red" nutrients were available, those flagged "amber" were chosen. In cases when only 2 nutrients were flagged "red" or "amber," a third 'green' (recommended) nutrient was given with a positive message to maintain nutrient intake.

Development of Decision Trees to Link Nutritional Intake Data to Feedback Messages

Sixteen dietary decision trees were manually developed to provide dietary feedback advice (Level 1 PN). All decision trees were developed to link nutrient intakes generated automatically by the Web-based Food4Me FFQ to a library of feedback messages. With the exception of the decision trees for carbohydrate, unsaturated fat, dietary fiber, and salt, all decision trees were generated to account for nutritional intake from both foods and supplements. Decision trees were based on the IOM gradation system with branches developed for "low," "recommended," and "high" intakes of each nutrient. The decision trees for SFA and salt were more complex and involved identifying the 2 main food groups contributing to the intake of the nutrient thereby enabling further personalization of dietary advice. For example, if a participant's diet were identified as high in SFAs, researchers would follow the decision tree to identify the 2 highest contributing food groups of 7 potential candidate groups.

All decision trees included branches linking to individual feedback messages. In total, the archive of dietary feedback messages consisted of 92 messages, which were developed using a variety of reputable sources including British Dietetic Association, Food Safety Authority Ireland, and British Nutrition Foundation [32-36]. For each decision tree, the feedback messages consisted of practical food-based tips to improve nutrient intake. An example of the decision tree for vitamin C and corresponding feedback messages are presented in Figure 2, with further examples of SFA and salt given in Multimedia Appendix 3. Protocols for the feedback system were standardized across the 7 countries, and all feedback messages were translated into the language of each recruitment country. To assess the utility and applicability of the decision trees and corresponding feedback messages, during the Food4Me study, a record was kept by researchers at all the 7 centers. These entries were amalgamated and categorized by nutrient.

Figure 2. Vitamin C decision tree and corresponding feedback messages.
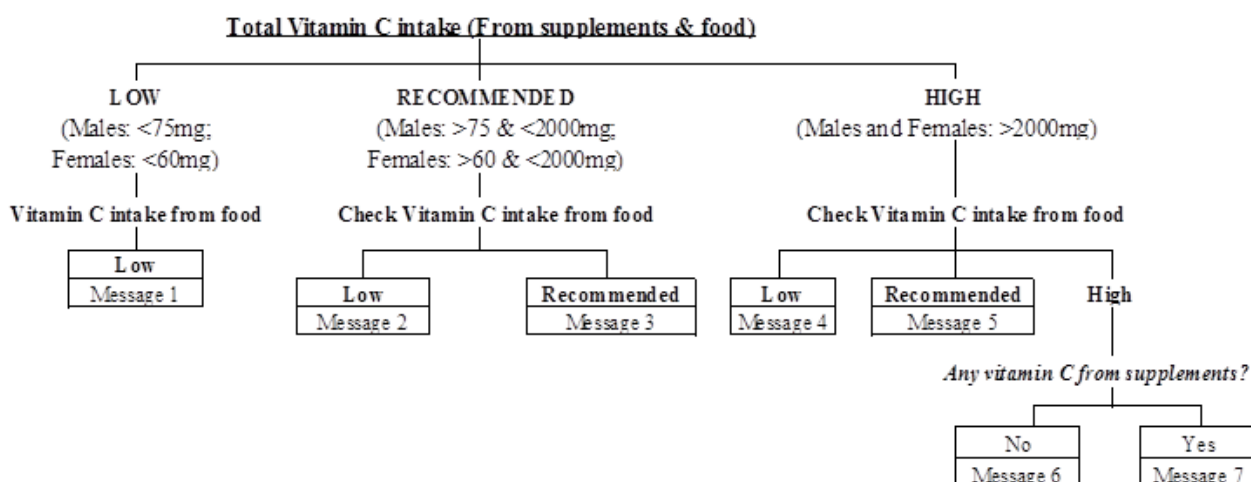
Message 1	How you can increase your vitamin C intake: • Eat more fruit and vegetables – aim for at least 5 a day • Try having a glass of fruit juice with breakfast • Berries are bursting with goodness – try adding them to cereals at breakfast, snack on them between meals or have them with low-fat natural yoghurt for dessert • Have peppers or tomatoes with salads or on sandwiches
Message 2	Your total vitamin C intake is within the recommended levels. This is a result of your consumption of supplements as your vitamin C intake from food is below the recommended levels. You should try to increase your intake of vitamin C: • Eat more fruit and vegetables • Try having a glass of fruit juice with breakfast • Berries are bursting with goodness – try adding them to cereals at breakfast, snack on them between meals or have them with low-fat natural yoghurt for dessert • Have peppers or tomatoes with salads or on sandwiches
Message 3	You are doing really well! Your vitamin C intake is within the recommended levels. It is important that you continue to maintain your healthy intake of this nutrient.
Message 4	Your total vitamin C intake is higher than recommended. This is a result of your consumption of supplements as your vitamin C intake from food is below the recommended levels. You should try to reduce the amount of supplements you take and aim to increase your intake of foods rich in vitamin C: • Eat more fruit and vegetables • Try having a glass of fruit juice with breakfast • Berries are bursting with goodness – try adding them to cereals at breakfast, snack on them between meals or have them with low-fat natural yoghurt for dessert • Have peppers or tomatoes with salads or on sandwiches
Message 5	Your total vitamin C intake is higher than recommended. As you are already meeting the dietary recommendations for vitamin C, your consumption of supplements is unnecessary.
Message 6	Your intake of vitamin C from the diet is exceeding the recommended levels. While it is important that you know that you are consuming quantities in excess of your requirements, having too much vitamin C is not associated with any adverse effects. You should however try to make sure you are consuming a variety of fruits and vegetables.
Message 7	Your total vitamin C intake is higher than recommended. As you are already exceeding your vitamin C requirements from your food intake, your intake of supplements is unnecessary. You should try to make sure you are consuming a variety of fruits and vegetables.

Development of Feedback Reports for Dietary Feedback Advice

Template feedback reports were developed to ensure that dietary advice was delivered in a consistent format across all countries.

Standard operating procedures were compiled and nutritionists or dietitians in all centers received training in composing the reports. All feedback reports began with a short message of encouragement from the researcher, which also highlighted the

main areas for improvement. This message was followed by 4 key sections:

- 1: How your diet compares to recommendations
- 2: Your physical characteristics
- 3: Your nutrient profile
- 4: Your personalized nutrition advice

Section 1 included a table comparing the participant's average number of portions of 5 food groups: *"fruit and vegetables," "whole grains," "dairy products," "oily fish,"* and *"red meat"* with guideline amounts. The guideline amounts were derived from amalgamating the national dietary advice in each of the 7 centers to create 1 common set of Food4Me dietary guidelines. Section 2 detailed the participant's height and weight and compared the participant's body mass index (BMI) and PA with recommendations. Section 3 graphically illustrated the participant's intake of each of the 17 nutrients as "good, no change recommended," "improvement recommended," and "improvement strongly recommended" on a gradation scale comparing intakes with IOM recommendations. Section 4 detailed personalized behavioral goals including a table listing the participant's 3 nutrient-related goals, along with dietary sources and the feedback message(s) from the corresponding decision tree. A list of all the food groups and nutrients for which personalized feedback was given can be found in Multimedia Appendix 4, with an example feedback report presented in Multimedia Appendix 5.

Automation of the Dietary Feedback

Nutrient intake analysis was automated for use in the Food4Me study, with the nutrient-related goals and feedback messages derived manually (using the priority system and decision trees) by Food4Me researchers using their own judgment to overrule the feedback system when appropriate. After completion of the Food4Me intervention study, the feedback system was automated by Creme Global (Dublin, Ireland). Therefore, a comparison of the automated and manual-based approaches was possible. The system was automated by firstly capturing the manual decision process for selecting priority nutrients, based on nutrient intake analysis, in a computer algorithm (written in the PHP programming language). Second, the decision trees for selection of feedback messages were encoded in table-like data structures and stored in a relational database (MySQL). Additional computer algorithms for traversing these data

structures were then developed (also using PHP), which enabled feedback messages to be generated automatically from nutrient intake analysis.

Analysis

The first part of the analysis examines the manual use of the feedback system during the Food4Me PoP study. Descriptive statistics were computed to describe the general characteristics of participants across each of the 7 centers using general linear model analysis with least significant difference post hoc. Nutrient-related goal selection frequency was examined across each of the 7 countries. All statistical analyses were conducted using IBM SPSS Statistics version 20, and $P<.05$ was considered statistically significant.

The second part of the analysis compares the dietary feedback provided by researchers during the Food4Me PoP study with feedback generated automatically by the computerized algorithms and messaging system for participants in Level 1 of the intervention. The level of agreement between the manual (researcher) and automated systems was assessed for both nutrient-related goal selection and feedback advice using baseline data for Level 1 participants, as illustrated in Figure 3. Agreement between the manual and automated systems, in selecting the 3 nutrient-related goals, was investigated by comparing each individual nutrient-related goal (1, 2, and 3) selected by both systems (manual vs automated) for each participant in the 7 countries. To evaluate the agreement for nutrient-related goal selection, we also examined whether the same 3 nutrients were selected as the 3 nutrient-related goals (in any order 1, 2, or 3) by both systems.

As outlined in Figure 3, to compare if the same feedback advice message(s) were derived from each decision tree by both systems, we selected only participants given the same nutrient-related goals (1, 2, and 3) by both the manual and automated systems and cross-compared the messages given by both systems for each nutrient-related goal (1, 2, and 3). Agreement was categorized as either "complete agreement," "complete disagreement," or "partial agreement" (data not shown). Overall agreement for feedback advice message was computed by summing the number of participants categorized into the 3 agreement groups for each nutrient-related goal across the countries.

Figure 3. Steps taken to compare the level of agreement between the manual and automated systems for nutrient-related goals and feedback advice. The agreement between the 2 systems in selecting the 3 nutrient-related goals was assessed for each participant. Nutrient-related goal agreement between the 2 systems was also evaluated for all 3 nutrient-related goals in random order. Where there was agreement between the 2 systems for nutrient goal selection, agreement in the feedback advice message(s) selected was then assessed between the two systems. In Figure 3, a represents complete agreement, 100% match between advice messages given by the manual and automated systems; b represents complete disagreement, 0% match between the advice messages given by the 2 systems; c represents partial agreement, some agreement, and some disagreement between the advice messages given by the 2 systems, partial agreement was only applicable to the SFA and salt decision trees as multiple advice messages were given in the feedback to participants.
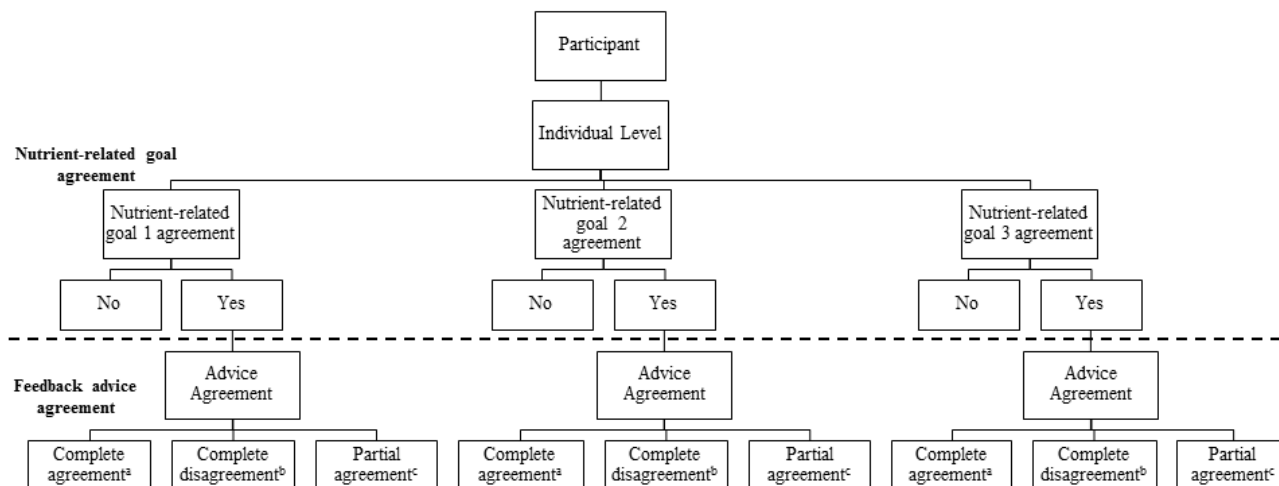
Results

The Study Population

A total of 414 participants, from the Level 1 group, across the 7 countries were available for inclusion in the analysis. After removing dropouts immediately after being randomized (n=41) and participants who did not have nutrient-related goals recorded at baseline (n=4), 369 participants were included in the analysis (Germany, n=52; Greece, n=51; Ireland, n=55; Netherlands, n=56; Poland, n=49; Spain, n=54; and the United Kingdom, n=52). There were no significant differences in age for men and women; however, self-reported BMI was significantly lower for women compared with men (P=.002). When comparing demographic characteristics across all the countries, no significant differences were observed for bodyweight or BMI; however, significant differences were observed for age (P=.02) and height (P<.001), as summarized in Table 1.

Table 1. Demographic characteristics of the study population in total and across countries[a].

Country	Demographic characteristics				
	Gender—female (%)	Age (y)	Height (m)	Weight (kg)	BMI (kg/m^2)
All countries (n=369)	57.5	39.7 ± 12.9	1.71 ± 0.10	74.19 ± 16.61	25.21 ± 5.02
Germany (n=52)	51.9	42.6 ± 14.9[b,d]	1.75 ± 0.09[b]	73.48 ± 13.55	23.95 ± 3.62
Greece (n=51)	58.8	38.1 ± 10.5[b,c]	1.69 ± 0.10[c]	76.02 ± 19.37	26.46 ± 6.52
Ireland (n=55)	56.4	38.9 ± 12.0[b,c,d]	1.70 ± 0.10[c]	72.90 ± 16.13	25.07 ± 4.76
Netherlands (n=56)	50.0	43.0 ± 15.4[d]	1.76 ± 0.10[b]	77.70 ± 16.29	25.08 ± 4.37
Poland (n=49)	71.4	36.2 ± 11.1[c]	1.69 ± 0.07[c]	71.46 ± 16.61	24.83 ± 4.99
Spain (n=54)	51.9	41.8 ± 11.0[b,d,e]	1.69 ± 0.10[c]	74.83 ± 16.94	26.09 ± 4.98
United Kingdom (n=52)	63.5	37.1 ± 13.4[c,e]	1.70 ± 0.76[c]	72.60 ± 17.05	24.94 ± 5.36

[a]Values are means ± standard deviations.

[b,c,d,e]Means with different superscripts denote significant differences (analysis of variance with least significant difference post hoc)

Manual Use and Efficacy of the Dietary Feedback System

The top 3 most frequently selected nutrient-related goals (1, 2, and 3) across each of the countries are summarized in Table 2. Saturated fatty acid was most frequently selected as nutrient-related goal 1 across all centers except for Spain where PUFA was most frequently selected as nutrient-related goal 1 in feedback given to 35% (19/54) of participants. Salt was most frequently selected as nutrient-related goal 2 in all 7 centers, with the exception of Germany, where both folate and salt were most frequently selected. As summarized in Table 2, greater variation was observed across the 7 countries for nutrient-related goal 3. As a result of the prioritization process, salt was most frequently selected as both nutrient-related goals 2 and 3, in both Spain and the United Kingdom.

The number of times each of the 17 nutrients was selected as a nutrient-related goal (1, 2, or 3) across each of the 7 countries is summarized in Table 3. Overall, SFA, salt, and dietary fiber were the top 3 most frequently targeted nutrient-related goals provided to 72%, 72%, and 39% of participants, respectively. Saturated fatty acid was the most frequently chosen nutrient-related goal overall for Germany, Ireland, and Poland. Salt was the most frequently selected nutrient-related goal overall for the Netherlands, Spain, and the United Kingdom, and in Greece, both dietary fiber and salt were most frequently selected.

A summary of the key issues encountered during the Food4Me study regarding the use and applicability of the decision trees and feedback advice messages is summarized in Table 4. Most of these issues were related to the SFA and salt decision trees, owing to their more complex design.

Table 2. Top 3 most frequently selected nutrient-related goals (1, 2, and 3) at baseline in the 7 countries[a].

Country	Nutrient-related goal 1	Nutrient-related goal 2	Nutrient-related goal 3
Germany (n=52)	SFA[b](54%)	Folate or salt (29%)	Carbohydrate or calcium (15%)
Greece (n=51)	SFA (41%)	Salt (39%)	Dietary fiber (41%)
Ireland (n=55)	SFA (56%)	Salt (53%)	Dietary fiber (15%)
Netherlands (n=56)	SFA (39%)	Salt (38%)	Dietary fiber 21%)
Poland (n=49)	SFA (53%)	Salt (39%)	Calcium (19%)
Spain (n=54)	PUFA[c](35%)	Salt (35%)	Salt (24%)
United Kingdom (n=52)	SFA (48%)	Salt (46%)	Salt (21%)

[a]Percentages indicate the percentage of participants who received the nutrient as the nutrient-related goal.

[b]SFA: saturated fatty acid.

[c]PUFA: polyunsaturated fatty acid.

Table 3. Number of participants having nutrient-related goals (1, 2, or 3) at baseline across the 7 countries[a].

Nutrient[a]	Germany (n=52)	Greece (n=51)	Ireland (n=55)	Netherlands (n=56)	Poland (n=49)	Spain (n=54)	United Kingdom (n=52)
Total fat	0	5	3	2	3	8	1
SFA[b]	42	27	47	43	38	30	37
MUFA[c]	6	2	3	13	11	3	12
PUFA[d]	6	16	8	3	3	21	13
Omega-3	8	6	2	6	0	2	4
Protein	2	0	2	4	2	0	1
Carbohydrate	9	1	4	2	2	1	3
Dietary fiber	14	31	20	19	14	24	20
Folate	23	23	8	13	23	20	10
Salt	31	31	46	47	31	35	45
Calcium	11	9	7	11	12	13	4
Iron	1	0	0	1	0	0	0
Vitamin C	2	1	3	2	4	2	1
Vitamin A	0	1	2	2	4	2	1
Vitamin B12	1	0	0	0	0	0	1
Total number of nutrient-related goals[e]	156	153	155	168	147	161[c]	153

[a]Riboflavin and thiamin were not given as nutrient-related goals in any of the 7 centers and are not presented in Table 3.

[b]SFA: saturated fatty acid.

[c]MUFA: monounsaturated fatty acid.

[d]PUFA: polyunsaturated fatty acid.

[e]Total number of nutrient-related goals in each country, calculated as number of participants (N) multiplied by 3. In Ireland, Spain, and the United Kingdom, total number of nutrient-related goals is less than this calculation as several participants were only given 2 nutrient-related goals in Ireland (n=10), Spain (n=1), the and United Kingdom (n=3).

Table 4. Issues with dietary decision trees or feedback messages.

Issue with decision trees or feedback messages	Decision trees affected
Algorithm did not capture all food groups contributing to the nutrient intake.	SFA[a]
Sometimes, the feedback messages are not specific to the participant's diet, that is, the participant's main sources of the nutrient may not have been identified in the feedback message.	SFA, salt, total fat
Feedback messages are repeated for different food groups.	Salt
Messages are not always relevant for vegetarians.	SFA, salt, omega 3

[a]SFA: saturated fatty acid.

Comparison of Advice Generated by Manual and Automated Systems

The level of agreement between the manual and automated systems for the selection of the 3 nutrient-related goals in all the countries is summarized in Table 5. Good agreement was observed between both methods in all the countries for nutrient-related goals 1 and 2 with average agreement of 92% and 87%, respectively. Agreement between the 2 methods ranged from 100% (the United Kingdom) to 82% (Greece) for nutrient-related goal 1 and from 98% (Spain) to 80% (the Netherlands) for nutrient-related goal 2. Lower agreement was observed between the manual and automated systems for nutrient-related goal 3, mean 63% across centers. For nutrient-related goal 3 selection, agreement was highest for Spain (85%) and lowest for Greece (45%). Agreement between the 2 systems for all 3 nutrient-related goals in random order ranged from 83% (Spain) to 47% (Greece), averaging 66% across all countries.

Having selected the priority nutrients, the next stage was the selection of feedback messages. The level of agreement between the manual and automated systems for the feedback advice

XSL·FO
RenderX

messages is summarized in Table 6. Complete agreement between the 2 systems for feedback advice ranged from 90% (Greece and Ireland) to 82% (Germany and Poland), with an average of 87% across the 7 countries.

Table 5. Level of agreement between manual and automated systems for baseline nutrient-related goals selection in all 7 countries[a].

Country	Nutrient-related goal 1	Nutrient-related goal 2	Nutrient-related goal 3	Random order agreement[b]
	n (%)	n (%)	n (%)	n (%)
Germany (n=52)	50 (96)	47 (90)	33 (63)	36 (69)
Greece (n=51)	41 (82)	42 (82)	23 (45)	24 (47)
Ireland (n=55)[c]	48 (87)	45 (82)	31 (56)	33 (60)
Netherlands (n=56)	50 (89)	45 (80)	34 (61)	36 (64)
Poland (n=49)	44 (90)	43 (88)	35 (71)	37 (76)
Spain (n=54)[c]	53 (98)	53 (98)	46 (85)	45 (83)
United Kingdom (n=52)[c]	52 (100)	48 (92)	31 (60)	33 (63)
Average	(92)	(87)	(63)	(66)

[a]Level of agreement between the manual and automated systems for each individual nutrient-related goal and for all 3 nutrient-related goals in random order.

[b]Agreement for all 3 nutrient-related goals in random order derived by calculating the agreement when the same 3 nutrients were selected as the 3 nutrient-related goals by both systems regardless if they were given as nutrient-related goal 1, nutrient-related goal 2, or nutrient-related goal 3.

[c]Lower agreement observed for nutrient related-goal 3 and overall agreement as a minority some participants received only 2 nutrient-related goals from the manual system in Ireland (n=10), Spain (n=1), and the United Kingdom (n=3).

Table 6. Level of agreement between manual and automated systems for baseline feedback advice message selection in all 7 countries.

Country	Complete agreement[a]	Complete disagreement[b]	Agreement plus disagreement[c]
	n (%)	n (%)	n (%)
Germany (n=130)	107 (82)	3 (2)	20 (16)
Greece (n=107)	95 (90)	0 (0)	12 (11)
Ireland (n=124)	112 (90)	3 (3)	9 (7)
Netherlands (n=129)	114 (88)	0 (0)	15 (12)
Poland (n=122)	100 (82)	1 (1)	21 (17)
Spain (n=152)	134 (88)	5 (3)	13 (9)
United Kingdom (n=131)	113 (86)	3 (2)	15 (11)
Average	(87)	(2)	(12)

[a]Complete agreement, 100% match between advice messages given by the manual and automated systems.

[b]Complete disagreement, 0% match between the advice messages given by the 2 systems.

[c]Partial agreement, some agreement, and some disagreement between the advice messages given by the 2 systems, partial agreement was only applicable to the SFA and salt decision trees as multiple advice messages were given in the feedback to participants.

Discussion

Principal Findings

This paper presents a novel system for providing consistent personalized dietary advice, automatically generated from dietary intake data submitted via the Internet by participants in a multi-country study. To our knowledge, this paper is one of the first to describe, in detail, the steps involved in developing a dietary feedback system to translate food and nutrient intake data into automatically generated personalized feedback advice and to compare the automatically generated advice with advice manually provided by nutrition researchers.

Development of the feedback system consisted of a 4-step process: designing a gradation system to categorize nutrient intakes, creating a priority system to enable 3 nutrients to be further selected as nutrient-related goals for particular emphasis in the feedback report, constructing decision trees to link nutrient intakes to feedback advice, and finally, developing feedback report templates. The findings of this paper demonstrate that automation of the feedback system is feasible and more superior compared with manual use of the system by result of removal of human error.

The feedback system described in this paper is unique as it was developed to provide personalized feedback on intakes of 5 food groups and 17 nutrients, 3 of which were prioritized and

XSL·FO
RenderX

selected as "nutrient-related goals" for subsequent targeted dietary advice. As the feedback system was developed to deliver relevant PN advice for all European adults, it was important that the system considered all nutrients for which changes in intake would be recommended to improve the health of significant proportions of the adult population. To address this requirement, the system that we developed included key food groups and a large number of nutrients so that feedback could be generated and personalized for a complete nutritional profile, rather than limited personalized information on one or several nutrient(s) as previously applied in other tailoring studies [20,37,38]. Providing extensive information, guaranteed that the dietary feedback would include personalized advice, which could be translated into food groups and nutrients of importance for a healthy diet (including fruit and vegetables, SFA, salt, fiber, whole grains, red meat and dairy products, omega-3 fatty acids, vitamin C, and iron) [39]. Several studies have briefly alluded to the application of a similar multistep process, linking nutrient intake data to feedback messages [1,13,21,24,37,40,41]. However, most of these studies have focused mostly on tailoring advice for selected nutrients or food groups only (eg, SFA, fruit and vegetable intakes) [1,6,13,16,21,24,37,42-44], with the exception of the advice generated for adolescents by Maes et al [40] which included fiber, vitamin C, calcium, iron, and fat and the reports developed by Kannan et al [23] that provided personalized feedback for 13 nutrients.

In addition to providing personalized food-based messages, feedback was also displayed graphically in reports by comparison of the individual's nutrient intakes with dietary recommendations, similar to many other studies [1,6,15,16,42,45]. Nutrient-related goal selection was relatively consistent across the countries, with SFA and salt being the most frequently assigned nutrient-related goals targeted, although some North–South gradient differences are observed. This result was expected given the prevalence of diets high in SFAs and salt across Europe [39,46,47].

Because the feedback system had not been tested previously in a multicenter cohort, the decision trees were used manually for the delivery of dietary advice within the intervention study. Manual use of the decision trees within the study combined with rigorous recording of the issues identified by the nutritionists or dietitians delivering the intervention facilitated evaluation of the advice and identification of aspects of the system, which could be improved. The decision trees were developed to comprise essential components such as contribution from supplements and food groups (SFA and salt decision trees only). To guarantee feedback advice is fully relevant and appropriately personalized at an individual level, suggested improvements include further expanding the decision trees to integrate specific food items and incorporate additional food groups (SFA and salt decision trees), for example, branches for consumption of commonly consumed food items high in SFAs (eg, ice cream and quiche) could be added to the SFA and total fat decision trees.

Automated technologies are increasingly used across many disciplines and have been shown to be as effective as human (face-to-face) systems for 3-month weight loss and exercise interventions [8,48]. Furthermore, Emerencia et al [49] recently developed an automated personalized system for schizophrenia patients providing advice similar to that given by clinicians. Although several studies have generated computer-automated dietary feedback, to our knowledge, this paper is among the first to examine whether it is possible to fully automate a dietary feedback system designed to provide PN advice to enhance healthy eating. We observed high agreement between the manual and automated systems, averaging at 92% and 87%, for nutrient-related goals 1 and 2, respectively. Good agreement was also observed for feedback advice selection averaging 87%. However, although our results show potential for future automation of dietary feedback tools, some disagreement was observed between the manual and automated systems.

Nutrient goal selection disagreement averaged 8%, 13%, and 37% across the 7 centers for nutrient-related goals 1, 2, and 3, respectively. Reasons for this disagreement included: researchers overruling the priority system (eg, if they thought it was inappropriate to give one of the nutrient-related goals selected using the priority system); researchers misreporting the nutrient-related goal in the report (eg, the correct nutrient-related goal was given in the report but misreported in the Internet); and researcher error in selecting the nutrient-related goal (eg, giving an adequate nutrient by mistake). The level of disagreement between the 2 systems was much greater for nutrient-related goal 3, especially for Greece, primarily, a result of overruling the priority system. Overruling did not always mean that the original selection was inappropriate, only that the researcher thought another option was more relevant. For feedback advice selection, complete disagreement was negligible (2%), but partial agreement was higher (12%), and was confined to the SFA and salt decision tree messages only. Owing to their more complex design, the feedback advice for these nutrients included 2-4 individual messages, which were more prone to researcher misreporting, error or overruling the decision trees to give additional messages. There was also a degree of ambiguity in the prioritization rules and the SFA and salt decision tree rules. For example, in relation to prioritization, no rule was set for selecting a positive third nutrient when only 2 risk nutrients were identified, thereby increasing the level of disagreement for nutrient goal selection between the manual and automated systems. Similarly, for both the SFA and salt decision trees, no rules were established for which food groups should be selected when several contributed equally to the nutrient intake.

Although we observed some disagreement between the manual and automated systems, overall, the agreement between the systems was excellent. Much of the disagreement we observed between the 2 systems would be reduced by automation of the feedback system, which would eliminate researcher error and researcher misreporting, and further clarification of the decision rules for both the prioritization process and SFA and salt decision trees. Furthermore, despite researchers overruling the priority nutrients, both the systems selected nutrients of high priority.

Technological innovations have evolved the delivery and management of health care through the increasing use of mobile phone health and computerized health care apps [50,51], with the personalized health field developing rapidly [52].

Consequently, our feedback system has potential to become a cost-effective approach to improve dietary behaviors at an individual level, which could be delivered by public or private health care providers. The present system was designed to provide advice on usual intake rather than actual intake; the challenge in the future will be to provide advice to consumers at the point of purchase or consumption.

Strengths and Limitations

The strengths of this study include the adequate sample size and multicenter recruitment in 7 European countries. In addition, to facilitate behavior change, the system was designed to incorporate numerous behavioral change techniques including goal setting, action planning, and barrier identification [11,25]. However, the system did not take into account motivation for change, and this could be a potential limitation. Feedback that is personalized to current dietary intake and stage of change may be considered as being more motivational and could have greater efficacy in promoting sustained dietary changes than advice personalized on dietary intake data only [14]. Another potential limitation of the study is that the feedback system was tested in a group of individuals who volunteered to receive PN

advice and may therefore not be applicable to other population groups. However, the dietary and anthropometric characteristics of participants in the Food4Me study were broadly similar to those of the wider population of European adults so that the study tested the utility of the system to select appropriate feedback for the likely most common dietary changes. In addition, we did not follow-up with participants regarding their opinions about the advice they received, for example, if they thought it was relevant to them, appropriate, or useful, and this would have been extremely useful information to have captured for the future progression of PN.

Conclusions

We developed a Web-based dietary feedback system that was capable of delivering consistent personalized dietary advice to adult European participants in the multicenter Food4Me study. Outcomes from comparison of the manual and automated feedback systems provide confidence that such an automated dietary feedback system can be developed and implemented across multiple countries with the potential to contribute to scalable and cost-effective interventions to improving dietary behaviors and health across large populations.

Authors' Contributions

The authors' responsibilities were as follows: HF, MCW, LB, and EG carried out the analyses and drafted the manuscript; CCM, HF, CBO, CW, CG, CFMM, ALM, RF, SNC, RSC, SK, LT, CPL, MG, AS, IMFG, MCW, ERG, LB, and JCM contributed to the standardized operating procedures for the study; EJD and RO contributed to the design and technical aspect of the Web-based server for automating the feedback system; CCM, SNC, RSC, HF, CW, CBO, CFMM, AM, RF, SK,CM, LT, CPL, MG, AS, MCW, and JCM conducted the intervention; YM, IT, CAD, ERG, LB, JAL, JAM, WHMS, HD, MJG, and JCM contributed to the research design of the Food4Me study. All authors contributed to a critical review of the manuscript during the writing process and approved the final version to be published.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Nutrient gradations.

[PDF File (Adobe PDF File), 597KB - jmir_v18i6e150_app1.pdf]

Multimedia Appendix 2

Level 1 priority nutrients and rationale.

[PDF File (Adobe PDF File), 294KB - jmir_v18i6e150_app2.pdf]

Multimedia Appendix 3

Example decision trees and messages.

[PDF File (Adobe PDF File), 244KB - jmir_v18i6e150_app3.pdf]

Multimedia Appendix 4

Nutrients and food groups for which personalized feedback was given.

XSL•FO
RenderX

[PDF File (Adobe PDF File), 250KB - jmir_v18i6e150_app4.pdf]

Multimedia Appendix 5

Example feedback report.

[PDF File (Adobe PDF File), 447KB - jmir_v18i6e150_app5.pdf]

References

1. Wright JL, Sherriff JL, Dhaliwal SS, Mamo John C L. Tailored, iterative, printed dietary feedback is as effective as group education in improving dietary behaviours: results from a randomised control trial in middle-aged adults with cardiovascular risk factors. Int J Behav Nutr Phys Act 2011;8:43 [FREE Full text] [doi: 10.1186/1479-5868-8-43] [Medline: 21595978]
2. Desroches S, Lapointe A, Ratté S, Gravel K, Légaré F, Turcotte S. Interventions to enhance adherence to dietary advice for preventing and managing chronic diseases in adults. Cochrane Database Syst Rev 2013;2:CD008722. [doi: 10.1002/14651858.CD008722.pub2] [Medline: 23450587]
3. McCarthy M. Higher sugar intake linked to raised risk of cardiovascular mortality, study finds. BMJ 2014;348:g1352. [Medline: 24501228]
4. Dhurandhar NV, Thomas D. The link between dietary sugar intake and cardiovascular disease mortality: an unresolved question. JAMA 2015 Mar 3;313(9):959-960. [doi: 10.1001/jama.2014.18267] [Medline: 25734737]
5. Swinburn B, Sacks G, Vandevijvere S, Kumanyika S, Lobstein T, Neal B, INFORMAS. INFORMAS (International Network for Food and Obesity/non-communicable diseases Research, Monitoring and Action Support): overview and key principles. Obes Rev 2013 Oct;14 Suppl 1:1-12. [doi: 10.1111/obr.12087] [Medline: 24074206]
6. Block G, Block T, Wakimoto P, Block CH. Demonstration of an E-mailed worksite nutrition intervention program. Prev Chronic Dis 2004 Oct;1(4):A06 [FREE Full text] [Medline: 15670437]
7. Probst YC, Tapsell LC. Overview of computerized dietary assessment programs for research and practice in nutrition education. J Nutr Educ Behav 2005;37(1):20-26. [Medline: 15745652]
8. 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. [doi: 10.1001/archinte.166.15.1620] [Medline: 16908795]
9. Juan W, Gerrior S, Hiza H. MyPyramid Tracker assesses food consumption, physical activity, and energy balance status interactively. J Nutr Educ Behav 2006;38(6 Suppl):S155-S157. [doi: 10.1016/j.jneb.2006.07.018] [Medline: 17116594]
10. van GL, van EP, Boon B, Borsboom G, Visscher T, Oenema A. Results from an online computer-tailored weight management intervention for overweight adults: randomized controlled trial. J Med Internet Res 2012;14(2):e44 [FREE Full text] [doi: 10.2196/jmir.1901] [Medline: 22417813]
11. Celis-Morales C, Lara J, Mathers JC. Personalising nutritional guidance for more effective behaviour change. Proc Nutr Soc 2015 May;74(2):130-138. [doi: 10.1017/S0029665114001633] [Medline: 25497396]
12. Anton SD, LeBlanc E, Allen HR, Karabetian C, Sacks F, Bray G, et al. Use of a computerized tracking system to monitor and provide feedback on dietary goals for calorie-restricted diets: the POUNDS LOST study. J Diabetes Sci Technol 2012 Sep;6(5):1216-1225 [FREE Full text] [Medline: 23063049]
13. Springvloet L, Lechner L, Oenema A. Planned development and evaluation protocol of two versions of a web-based computer-tailored nutrition education intervention aimed at adults, including cognitive and environmental feedback. BMC Public Health 2014;14:47 [FREE Full text] [doi: 10.1186/1471-2458-14-47] [Medline: 24438381]
14. Forster H, Walsh MC, Gibney MJ, Brennan L, Gibney ER. Personalised nutrition: the role of new dietary assessment methods. Proc Nutr Soc 2016 Feb;75(1):96-105. [doi: 10.1017/S0029665115002086] [Medline: 26032731]
15. Ezendam Nicole P M, Brug J, Oenema A. Evaluation of the Web-based computer-tailored FATaintPHAT intervention to promote energy balance among adolescents: results from a school cluster randomized trial. Arch Pediatr Adolesc Med 2012 Mar;166(3):248-255. [doi: 10.1001/archpediatrics.2011.204] [Medline: 22064878]
16. Kroeze W, Oenema A, Campbell M, Brug J. The efficacy of Web-based and print-delivered computer-tailored interventions to reduce fat intake: results of a randomized, controlled trial. J Nutr Educ Behav 2008;40(4):226-236. [doi: 10.1016/j.jneb.2007.09.008] [Medline: 18565463]
17. Lustria Mia Liza A, Noar SM, Cortese J, Van Stee Stephanie K, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. J Health Commun 2013;18(9):1039-1069. [doi: 10.1080/10810730.2013.768727] [Medline: 23750972]
18. Moore TJ, Alsabeeh N, Apovian CM, Murphy MC, Coffman GA, Cullum-Dugan D, et al. Weight, blood pressure, and dietary benefits after 12 months of a Web-based Nutrition Education Program (DASH for health): longitudinal observational study. J Med Internet Res 2008;10(4):e52 [FREE Full text] [doi: 10.2196/jmir.1114] [Medline: 19073541]
19. Patrick K, Calfas KJ, Norman GJ, Rosenberg D, Zabinski MF, Sallis JF, et al. Outcomes of a 12-month web-based intervention for overweight and obese men. Ann Behav Med 2011 Dec;42(3):391-401. [doi: 10.1007/s12160-011-9296-7] [Medline: 21822750]

XSL•FO
RenderX

20. Sternfeld B, Block C, Quesenberry CP, Block TJ, Husson G, Norris JC, et al. Improving diet and physical activity with ALIVE: a worksite randomized trial. Am J Prev Med 2009 Jun;36(6):475-483. [doi: 10.1016/j.amepre.2009.01.036] [Medline: 19460655]

21. Oenema A, Tan F, Brug J. Short-term efficacy of a web-based computer-tailored nutrition intervention: main effects and mediators. Ann Behav Med 2005 Feb;29(1):54-63. [doi: 10.1207/s15324796abm2901_8] [Medline: 15677301]

22. Illner A, Freisling H, Boeing H, Huybrechts I, Crispim SP, Slimani N. Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. Int J Epidemiol 2012 Aug;41(4):1187-1203 [FREE Full text] [doi: 10.1093/ije/dys105] [Medline: 22933652]

23. Kannan S, Schulz A, Israel B, Ayra I, Weir S, Dvonch TJ, et al. A community-based participatory approach to personalized, computer-generated nutrition feedback reports: the healthy environments partnership. Prog Community Health Partnersh 2008;2(1):41-53 [FREE Full text] [doi: 10.1353/cpr.2008.0004] [Medline: 19337572]

24. Brug J. Dutch research into the development and impact of computer-tailored nutrition education. Eur J Clin Nutr 1999 May;53 Suppl 2:S78-S82. [Medline: 10406443]

25. Celis-Morales C, Livingstone KM, Marsaux Cyril F M, Forster H, O'Donovan CB, Woolhead C, Matthews John N S, Saris Wim H M, et al. Design and baseline characteristics of the Food4Me study: a web-based randomised controlled trial of personalised nutrition in seven European countries. Genes Nutr 2015 Jan;10(1):450 [FREE Full text] [doi: 10.1007/s12263-014-0450-2] [Medline: 25491748]

26. Forster H, Fallaize R, Gallagher C, O'Donovan CB, Woolhead C, Walsh MC, et al. Online dietary intake estimation: the Food4Me food frequency questionnaire. J Med Internet Res 2014;16(6):e150 [FREE Full text] [doi: 10.2196/jmir.3105] [Medline: 24911957]

27. Fallaize R, Forster H, Macready AL, Walsh MC, Mathers JC, Brennan L, et al. Online dietary intake estimation: reproducibility and validity of the Food4Me food frequency questionnaire against a 4-day weighed food record. J Med Internet Res 2014;16(8):e190 [FREE Full text] [doi: 10.2196/jmir.3355] [Medline: 25113936]

28. Institute of Medicine. Dietary Reference Intakes for energy, Carbohydrate, Fibre, Fat, Fatty acids, Cholesterol, Protein, and Amino acids. 2005 URL: http://www.nap.edu/openbook.php?isbn=0309085373 [accessed 2016-01-25] [WebCite Cache ID 6enq6mN89]

29. Institute of Medicine. Dietary Reference Intakes Tables and Applications. 2011 URL: http://www.iom.edu/Activities/Nutrition/SummaryDRIs/DRI-Tables.aspx [accessed 2012-01-18] [WebCite Cache ID 6enoKWKJw]

30. World Health Organisation. Protein and Amino acid requirements in Human Nutrition. 2007 URL: http://apps.who.int/iris/bitstream/10665/43411/1/WHO_TRS_935_eng.pdf [accessed 2012-01-18] [WebCite Cache ID 6enoRIOnS]

31. World Health Organisation. Food and Agriculture Organisation of the United Nations (FAO) report of an expert consultation on fats and fatty acids in human nutrition. 2010. URL: http://www.who.int/nutrition/publications/nutrientrequirements/fatsandfattyacids_humannutrition/en/ [accessed 2012-01-20] [WebCite Cache ID 6enoXSpNi]

32. Public Health Agency. Enjoy Healthy Eating. 2012. URL: http://www.enjoyhealthyeating.info/nutrition/adults [accessed 2016-01-25] [WebCite Cache ID 6enogDfPj]

33. British Dietetic Association. Food Fact Sheets. 2013. URL: https://www.bda.uk.com/foodfacts/home [accessed 2016-01-25] [WebCite Cache ID 6enoqBvPV]

34. British Dietetic Association. Weight Wise Eating well what?s a healthy diet?. 2013. URL: http://www.bdaweightwise.com/eating/eating_diet3.html [accessed 2016-01-25] [WebCite Cache ID 6enox8ZRg]

35. British Nutrition Foundation. Healthy eating. 2013. URL: https://www.nutrition.org.uk/healthyliving/healthyeating.html [accessed 2016-01-25] [WebCite Cache ID 6enp5Zq4c]

36. Food Saftey Authority Ireland. Scientific Recommendations for Healthy Eating Guidelines in Ireland. 2011. URL: https://www.fsai.ie/science_and_health/healthy_eating.html [accessed 2016-01-25] [WebCite Cache ID 6enpE0UAs]

37. Ambeba EJ, Ye L, Sereika SM, Styn MA, Acharya SD, Sevick MA, et al. The use of mHealth to deliver tailored messages reduces reported energy and fat intake. J Cardiovasc Nurs 2015;30(1):35-43. [doi: 10.1097/JCN.0000000000000120] [Medline: 24434827]

38. Springvloet L, Lechner L, de VH, Candel Math J J M, Oenema A. Short- and medium-term efficacy of a Web-based computer-tailored nutrition education intervention for adults including cognitive and environmental feedback: randomized controlled trial. J Med Internet Res 2015;17(1):e23 [FREE Full text] [doi: 10.2196/jmir.3837] [Medline: 25599828]

39. Capacci S, Mazzocchi M, Shankar B, Macias JB, Verbeke W, Pérez-Cueto Federico J A, et al. Policies to promote healthy eating in Europe: a structured review of policies and their effectiveness. Nutr Rev 2012 Mar;70(3):188-200. [doi: 10.1111/j.1753-4887.2011.00442.x] [Medline: 22364161]

40. Maes L, Cook TL, Ottovaere C, Matthijs C, Moreno LA, Kersting M, et al. Pilot evaluation of the HELENA (Healthy Lifestyle in Europe by Nutrition in Adolescence) Food-O-Meter, a computer-tailored nutrition advice for adolescents: a study in six European cities. Public Health Nutr 2011 Jul;14(7):1292-1302. [doi: 10.1017/S1368980010003563] [Medline: 21281542]

41. Kerr DA, Pollard CM, Howat P, Delp EJ, Pickering M, Kerr KR, et al. Connecting Health and Technology (CHAT): protocol of a randomized controlled trial to improve nutrition behaviours using mobile devices and tailored text messaging in young adults. BMC Public Health 2012;12:477 [FREE Full text] [doi: 10.1186/1471-2458-12-477] [Medline: 22726532]

42.   Robroek Suzan J W, Bredt FJ, Burdorf A. The (cost-)effectiveness of an individually tailored long-term worksite health promotion programme on physical activity and nutrition: design of a pragmatic cluster randomised controlled trial. BMC Public Health 2007;7:259 [FREE Full text] [doi: 10.1186/1471-2458-7-259] [Medline: 17888161]

43.   Smeets T, Kremers S P J, Brug J, de VH. Effects of tailored feedback on multiple health behaviors. Ann Behav Med 2007 Apr;33(2):117-123. [doi: 10.1080/08836610701307801] [Medline: 17447863]

44.   Raats MM, Sparks P, Geekie MA, Shepherd R. The effects of providing personalized dietary feedback. A semi-computerized approach. Patient Educ Couns 1999 Jun;37(2):177-189. [Medline: 14528544]

45.   Oenema A, Brug J. Feedback strategies to raise awareness of personal dietary intake: results of a randomized controlled trial. Prev Med 2003 Apr;36(4):429-439. [Medline: 12649051]

46.   Brown IJ, Tzoulaki I, Candeias V, Elliott P. Salt intakes around the world: implications for public health. Int J Epidemiol 2009 Jun;38(3):791-813 [FREE Full text] [doi: 10.1093/ije/dyp139] [Medline: 19351697]

47.   Linseisen J, Welch AA, Ocké M, Amiano P, Agnoli C, Ferrari P, van der Schouw Y T, et al. Dietary fat intake in the European Prospective Investigation into Cancer and Nutrition: results from the 24-h dietary recalls. Eur J Clin Nutr 2009 Nov;63 Suppl 4:S61-S80. [doi: 10.1038/ejcn.2009.75] [Medline: 19888281]

48.   King AC, Hekler EB, Castro CM, Buman MP, Marcus BH, Friedman RH, et al. Exercise advice by humans versus computers: maintenance effects at 18 months. Health Psychol 2014 Feb;33(2):192-196. [doi: 10.1037/a0030646] [Medline: 23421896]

49.   Emerencia A, van der Krieke Lian, Sytema S, Petkov N, Aiello M. Generating personalized advice for schizophrenia patients. Artif Intell Med 2013 May;58(1):23-36. [doi: 10.1016/j.artmed.2013.01.002] [Medline: 23419698]

50.   Meier CA, Fitzgerald MC, Smith JM. eHealth: extending, enhancing, and evolving health care. Annu Rev Biomed Eng 2013;15:359-382. [doi: 10.1146/annurev-bioeng-071812-152350] [Medline: 23683088]

51.   Weinstein RS, Lopez AM, Joseph BA, Erps KA, Holcomb M, Barker GP, et al. Telemedicine, telehealth, and mobile health applications that work: opportunities and barriers. Am J Med 2014 Mar;127(3):183-187. [doi: 10.1016/j.amjmed.2013.09.032] [Medline: 24384059]

52.   Clifton L, Clifton DA, Pimentel Marco A F, Watkinson PJ, Tarassenko L. Gaussian processes for personalized e-health monitoring with wearable sensors. IEEE Trans Biomed Eng 2013 Jan;60(1):193-197. [doi: 10.1109/TBME.2012.2208459] [Medline: 23268532]

Abbreviations

BMI: body mass index
EAR: estimated average requirement
FFQ: food frequency questionnaire
IOM: Institute of Medicine
NCD: noncommunicable diseases
PA: physical activity
PN: personalized nutrition
PoP: Proof-of-Principle
PUFA: polyunsaturated fatty acids
SFA: saturated fatty acid
UL: upper limit

XSL•FO

RenderX

Original Paper

Mining Health App Data to Find More and Less Successful Weight Loss Subgroups

Katrina J Serrano[1*], PhD; Mandi Yu[1*], PhD; Kisha I Coa[2*], PhD, MPH; Linda M Collins[3*], PhD; Audie A Atienza[2*], PhD

[1]National Cancer Institute, Bethesda, MD, United States

[2]ICF International, Rockville, MD, United States

[3]Pennsylvania State University, State College, PA, United States

[*]all authors contributed equally

Corresponding Author:
Katrina J Serrano, PhD
National Cancer Institute
9609 Medical Center Dr.
Bethesda, MD, 20892
United States
Phone: 1 2402766654
Fax: 1 2402767907
Email: katrina.serrano@nih.gov

Abstract

Background: More than half of all smartphone app downloads involve weight, diet, and exercise. If successful, these lifestyle apps may have far-reaching effects for disease prevention and health cost-savings, but few researchers have analyzed data from these apps.

Objective: The purposes of this study were to analyze data from a commercial health app (Lose It!) in order to identify successful weight loss subgroups via exploratory analyses and to verify the stability of the results.

Methods: Cross-sectional, de-identified data from Lose It! were analyzed. This dataset (n=12,427,196) was randomly split into 24 subsamples, and this study used 3 subsamples (combined n=972,687). Classification and regression tree methods were used to explore groupings of weight loss with one subsample, with descriptive analyses to examine other group characteristics. Data mining validation methods were conducted with 2 additional subsamples.

Results: In subsample 1, 14.96% of users lost 5% or more of their starting body weight. Classification and regression tree analysis identified 3 distinct subgroups: "the occasional users" had the lowest proportion (4.87%) of individuals who successfully lost weight; "the basic users" had 37.61% weight loss success; and "the power users" achieved the highest percentage of weight loss success at 72.70%. Behavioral factors delineated the subgroups, though app-related behavioral characteristics further distinguished them. Results were replicated in further analyses with separate subsamples.

Conclusions: This study demonstrates that distinct subgroups can be identified in "messy" commercial app data and the identified subgroups can be replicated in independent samples. Behavioral factors and use of custom app features characterized the subgroups. Targeting and tailoring information to particular subgroups could enhance weight loss success. Future studies should replicate data mining analyses to increase methodology rigor.

KEYWORDS

weight loss; mobile health; mobile app; data mining; classification

Introduction

Smartphone ownership among American adults has increased from 35% in 2011 to 68% in 2015 [1]. This increase has coincided with the proliferation of smartphone apps, and 19% of all app downloads are related to health, with more than half of them involving weight, diet, and exercise [2]. This provides new opportunities to deliver interventions for health behavior change and weight loss in the United States where obesity rates have remained high [3].

Although apps show great promise for helping individuals lose weight and manage lifestyle habits [4-6], evidence to support the impact of commercial apps on health behavior and weight loss is still lacking. This may be due to the lack of evidence-based weight loss principles in currently available apps [7]. But given the popularity of these apps, the potential implications are far-reaching, not only in terms of disease prevention (eg, diabetes, cardiovascular diseases, cancer) but also in cost-savings [8-11].

Data that are collected from commercial health apps are often not collected with scientific research in mind. However, these apps can reach millions of users. If analyzed with rigorous scientific methods, the potentially rich data collected from these apps may offer important insights into how behavior change occurs in naturalistic settings among large segments of the population. Exploratory analyses, such as data mining methods, that can be used to examine existing health data are not new [11-13], but they have rarely been used to examine health data collected from commercial apps.

Furthermore, scientific methods to examine the reliability and robustness of exploratory analyses (ie, data mining validation methods) have also been available for some time [14,15], but have not been used with health app data. With millions of individuals using commercial health apps, opportunities now exist for both exploratory data mining and data mining validation methods to occur in rapid succession. Data mining validation methods increase the scientific rigor of exploratory approaches by testing whether initial findings are stable.

To our knowledge, no studies have explored the effectiveness of a weight loss commercial app AND evaluated the reliability of the exploratory findings. The purposes of this study were to (1) assess the prevalence of weight loss among overweight and obese adults from data gathered by a commercial app, (2) identify successful weight loss subgroups and their characteristics using exploratory data mining techniques, and (3) examine the reliability of the identified subgroups using independent samples.

Methods

Dataset

We analyzed a subset of cross-sectional, de-identified data (n=12,427,196), which were obtained directly from Lose It! (FitNow Inc., Boston, MA, USA). Data were made available to researchers at the National Cancer Institute for research purposes only. Lose It!—launched in 2008—is a weight loss app that is available through both iOS and Android app markets, as well as through the Web. Lose It! (henceforth, called the app) provides users with tracking tools (eg, barcode scanners); connections with other devices and apps (eg, Fitbit, RunKeeper); motivation and support (eg, connection with friends); and nutrition feedback (eg, system-generated reports comparing a user's food log with the US Department of Agriculture's MyPlate recommendations).

In the app, a user creates an account and a weight loss plan based on one's height, weight, exercise level, target weight goal, and desired weekly weight loss. The app then uses all this

information to calculate an estimated calories budget that is intended to produce the energy deficit required to meet one's weight loss goal. The weight loss plan consists of logging one's diet, exercise, and weight through either self-report or a synced device (eg, WiFi-connected body scales). The app offers motivation and support tools by allowing users to identify friends and share progress and information with them. Users can also participate in groups designed to motivate users; for example, one featured group—"We're all in this together!"—is described as "a group for people looking to give motivation and people looking to get motivation."

The data analyzed were from users who had the app during the years of 2008-2014. Data provided for analysis were from the app's metadata reporting database, which is used to power the app and provides a general summary of user activity. Thus, the data analyzed were cross-sectional in nature. The dataset included the following information: age at setup of the account, gender, height, body weight, body mass index (BMI), desired goal weight, desired weekly weight loss, number of days logged in for food and exercise, number of exercise calories burned, number of calories consumed, number of times weighed in, number of days active, date of last activity, devices and apps connected to a user's account, type of operating system used, number of friends and groups on the app, number of challenges users participated in, number of customized goals, foods, recipes, and exercises users entered, and app-specific options (eg, has a picture, uses reminders). Weight and health behavior data were self-reported, whereas technical-related data (eg, type of operating system used, app-specific options) were from the app's database. More time-intensive longitudinal data for the full sample of users between 2008 and 2014 were not readily available at the time of analyses.

Data cleaning was required before analyses, which included removing any duplicate records, placing valid ranges for each variable, and distinguishing between missing versus invalid data. There were 63,641 duplicates that were deleted. These users had the exact same information for all weight, health, and technical-related variables. We were left with a total sample of 12,363,555. Analyses with this entire sample proved to be challenging and required more computing memory than typically offered by a single computer. Therefore, for computing management and efficiency, this dataset was randomly split into 24 subsamples, each with a sample size of approximately 500,000. This study used 3 subsamples and excluded the following in each subsample: (1) participants who reported being less than 18 years or greater than 70 years at setup age—older adults (65 years and older) are less likely to use health-related smartphone apps [2], so to be more conservative, we chose 70 years as the upper age range; (2) participants who reported being younger than 18 years at the date of last activity; (3) participants who were underweight and of normal weight, BMI ≤24.9; and (4) participants with weight and weight loss values that were out of range; for example, we defined minimum weight values that exceeded start weight values as out of range.

The outcome of interest was weight loss, defined for the purpose of this study as losing 5% or more of a user's starting body weight, which has been shown to lead to beneficial health effects [16-18]. This was calculated by subtracting 5% of a user's

starting weight from a user's minimum weight. If this number was less than or equal to zero, then weight loss was categorized as yes, all others were categorized as no. The following predictors were included in the analyses: age, gender, number of weigh ins, target weight, weekly weight loss goal, start weight, start BMI, food and exercise days logged, average food and exercise calories logged, days active on the app, age at set up of the app, type of device or app used, type of operating system used, number of friends, number of groups, number of challenges, use of reminders, customized goals, customized recipes, customized exercises, and app-specific options.

Statistical analysis

Classification and regression tree (CART) analysis was conducted in subsample 1 (hereafter, known as the training sample). CART methods have been increasingly applied to health behavior research for exploratory purposes [19-23]. CART analysis is a type of decision tree methodology, also called recursive partitioning, that is useful for constructing prediction models from data [19,20,24-26]. CART uses nonparametric statistics to identify mutually exclusive and exhaustive subgroups of individuals who share common characteristics that influence the dependent variable of interest. The CART procedure uses a preselected splitting criterion to assess all possible independent variables and chooses a variable (ie, splitting variable) that results in binary groups that are the most different with regard to the dependent variable. The splitting criterion used was the Gini index of diversity [25], which selects the split that maximizes the reduction in impurity or diversity of a node, thereby reducing the error in classification [19,25].

CART methods have several advantages over more traditional approaches, such as logistic regression. Because CART is inherently nonparametric, no assumptions are made about the underlying distribution of the data. Thus, it can handle highly skewed distributions or even extreme scores or outliers [19,20,26]. CART also has sophisticated methods for handling missing data, and missing data are considered for each variable at each split point. If data are missing at a particular split point, surrogate variables that contain similar information to the primary splitter are used [27,28]. This is also an important consideration given the missing data typically seen in commercial health app data.

The CART analysis was conducted in R (version 3.1.3), using the package rpart. The default settings for rpart were used, and these parameters have been recommended by Breiman and colleagues [25]. More details about this package are provided elsewhere [28]. We then created mutually exclusive subgroups in the training sample based on the CART results. Descriptive analyses were conducted in SAS (version 9.3, SAS Institute, Inc., Cary, NC, USA) with the training sample to determine whether additional factors were uniquely associated with the various subgroups. Due to the large sample size, we were dubious of interpreting the P values; therefore, significance was determined by the unique variance explained by the predictor variables (using R^2 or Cramer's V). As a rule of thumb, the proportion of variance accounted for by the predictor variable had to be at least 1%.

The CART model predictions identified from the training sample were then evaluated with subsample 2 (hereafter, known as data mining validation sample 1) to examine the robustness of the model. The area under the receiver-operating characteristic curve (AUC) was used to evaluate the accuracy of the classification tree with data mining validation sample 1. Further evaluation was conducted with subsample 3 (hereafter, known as data mining validation sample 2), and the AUC was also obtained with this subsample. The AUC analyses were conducted in R (version 3.1.3), using the package pROC. More details about this package are provided elsewhere [29]. The annotated code regarding these analyses can be found here: https://github.com/kayserra/sample_code. For exploratory purposes, we also applied CART methods with data mining validation sample 2. We varied the default settings for the complexity parameter (ie, a criterion that takes into account the consequences of misclassification) to 0.001 versus 0.01 and the minimum number of observations in a node to compute a split as well as the terminal node to 3000 (1% of the sample) versus the default of 20 and 7, respectively.

Results

Analytic sample

Data cleaning and exclusion criteria applied to the 3 subsamples resulted in the following analytic samples: n=324,649 for subsample 1, n=324,063 for subsample 2, and n=323,975 for subsample 3 (data flow chart shown in Figure 1).

Figure 1. Data flow chart.
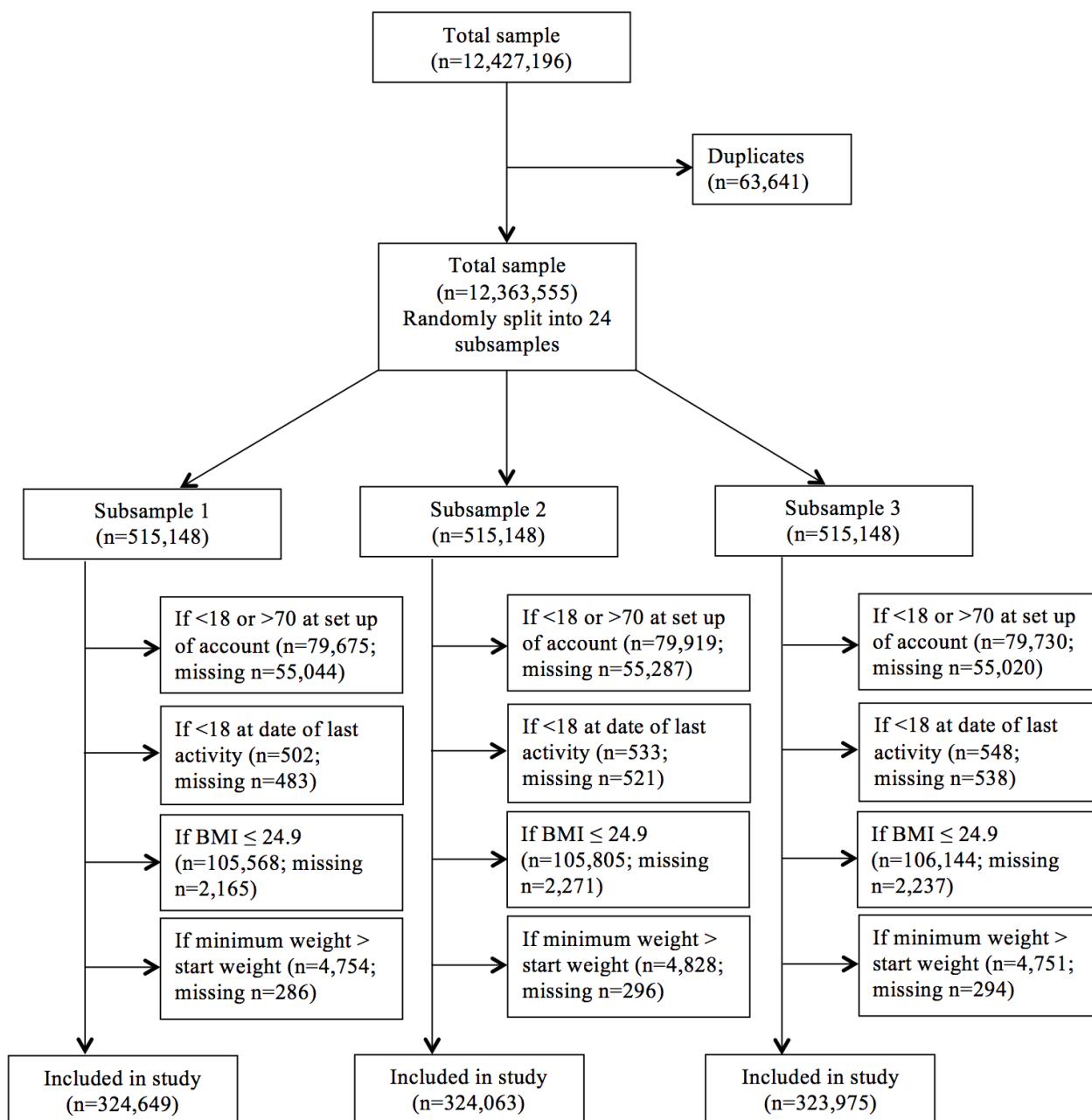
Statistical analysis

The CART model is displayed in Figure 2. As shown in the figure, 14.96% (48,562) of the training sample successfully lost weight. The CART analysis identified 3 distinct subgroups that we labeled for descriptive purposes: "the occasional users," "the basic users," and "the power users."

Although descriptive names are given for each subgroup, to more fully understand and interpret the subgroups, a set of additional characteristics were further examined. Results for the descriptive analyses that examined additional unique characteristics among the subgroups are displayed in Table 1.

XSL•FO
RenderX

Figure 2. Classification and regression tree for identifying successful weight loss subgroups with the training sample (n=324,649).
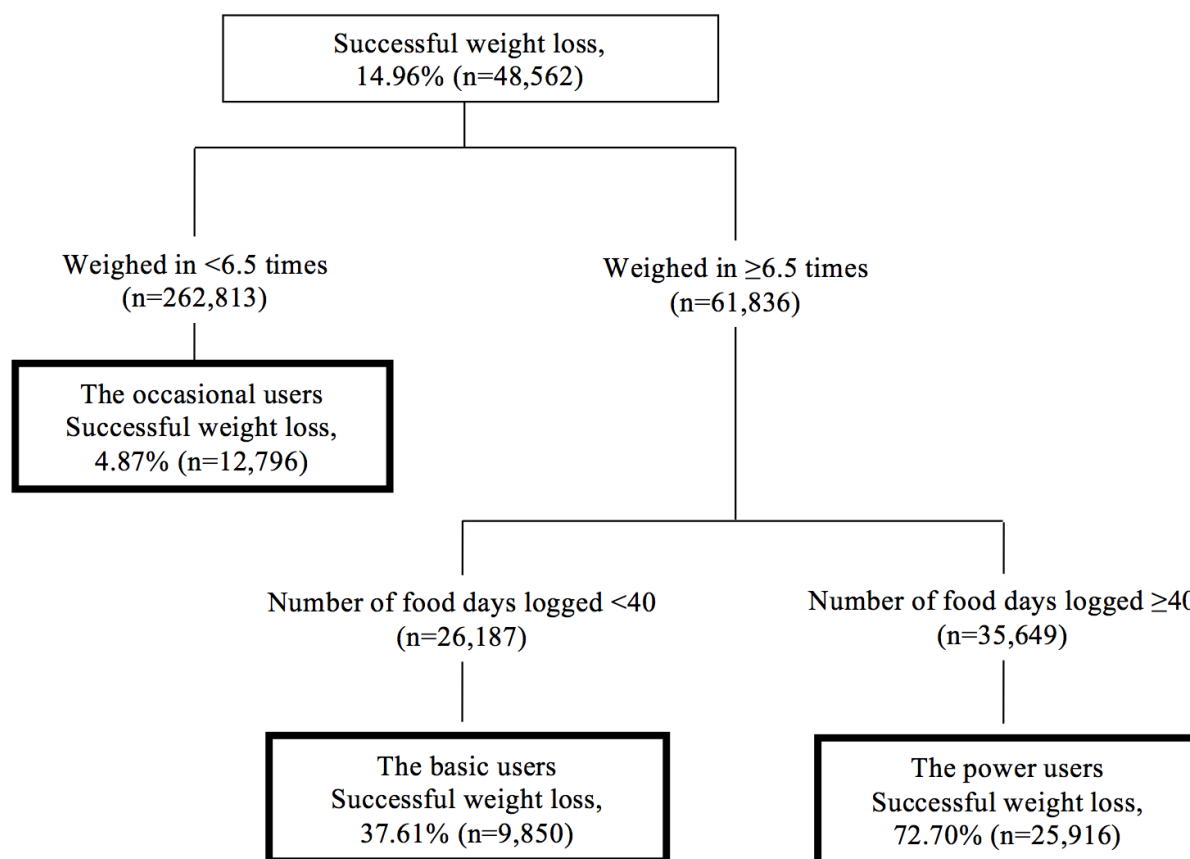
Table 1. Additional characteristics of identified successful weight loss subgroups with the training sample (n=324,649).

	The occasional users[a]	The basic users[b]	The power users[c]	Cramer's V	R^2	P
	% or mean (standard deviation)	% or mean (standard deviation)	% or mean (standard deviation)			
Demographics						
Female	74.32%	71.68%	63.53%	0.1049		<.001
Age (at set up of account)	34.5 (12.0)	35.4 (11.3)	39.0 (12.1)		0.0297	<.001
Start weight	212.0 (50.8)	211.3 (47.3)	211.2 (47.4)		0.0001	.274
Start BMI[d]	33.9 (7.0)	33.6 (6.6)	33.0 (6.4)		0.0034	<.001
Days active on the app	23.5 (46.0)	21.9 (10.4)	168.3 (174.7)		0.2383	<.001
Health behaviors						
Exercise days logged	9.8 (29.0)	9.0 (7.7)	80.5 (112.7)		0.1522	<.001
Exercise calories logged	39081969.7 (2931936775.4)	3799.2 (4979.1)	7753953.9 (1242356057.6)		0.0001	.169
Food calories logged	7844318.9 (884021907.5)	1040215.2 (100758647.6)	11818596.1 (1075871163.0)		0.0000	.602
Goal weight	160.5 (33.4)	161.8 (32.4)	166.2 (32.7)		0.0062	<.001
Goal plan[e]	1.7 (0.4)	1.8 (0.4)	1.6 (0.5)		0.0139	<.001
App behaviors						
iPhone users (% yes)[f]	71.59%	72.84%	77.73%	0.0653		<.001
Android users (% yes)[f]	29.40%	31.60%	30.88%	0.0171		<.001
Web users (% yes)[f]	4.01%	3.84%	2.94%	0.0277		<.001
One or more devices/apps linked with app (eg, Fitbit) (% yes)	3.70%	7.82%	14.00%	0.1487		<.001
Has friends on the app (% yes)	18.01%	27.37%	43.44%	0.2356		<.001
Number of friends on the app	0.3 (1.3)	0.6 (2.2)	2.1 (14.4)		0.0062	<.001
Is part of a group on the app (% yes)	1.41%	3.21%	5.45%	0.0894		<.001
Number of groups on the app	0.0 (0.3)	0.1 (0.4)	0.1 (1.1)		0.0039	<.001
Has been an administrator of a challenge (% yes)	0.02%	0.05%	0.32%	0.0332		<.001
Number of challenges participated in	0.0 (0.0)	0.0 (0.0)	0.0 (0.2)		0.0007	<.001
Number of customized goals entered	0.0 (0.4)	0.1 (0.7)	0.3 (1.4)		0.0127	<.001
Number of customized foods entered	5.9 (16.6)	7.7 (13.3)	43.9 (81.4)		0.0866	<.001
Number of customized recipes entered	0.4 (2.2)	0.5 (2.0)	4.3 (13.4)		0.0373	<.001
Number of customized exercises entered	0.5 (8.0)	0.6 (2.7)	3.1 (19.9)		0.0071	<.001
Uses app reminders (% yes)	5.97%	8.30%	14.23%	0.1189		<.001
Has a picture (% yes)	9.60%	15.54%	25.70%	0.1780		<.001

XSL•FO
RenderX

	The occasional users[a]	The basic users[b]	The power users[c]	Cramer's V	R^2	P
	% or mean (standard deviation)	% or mean (standard deviation)	% or mean (standard deviation)			
Uses email reports (% yes)	1.45%	2.88%	6.22%	0.1048		<.001

[a]4.87% achieved weight loss success (n=12,796).

[b]37.61% achieved weight loss success (n=9,850).

[c]72.70% achieved weight loss success (n=25,916).

[d]BMI, body mass index.

[e]Desired weekly weight loss (0-2 lbs).

[f]Users can download and access the app on multiple platforms and devices.

The occasional users achieved the lowest percentage of weight loss success (4.87%), and these users weighed in on the app <6.5 times. Approximately 37.61% of the basic users achieved at least 5% weight loss, and these individuals weighed in at least 6.5 times and logged in food <40 days. The power users had the highest percentage of weight loss success (72.70%) and consisted of individuals who weighed in at least 6.5 times and logged in food ≥40 days.

Compared with the other subgroups identified, the power users had more men (36.47%) than the occasional users or the basic users, and they were more active with the app (about 168 days). They also logged in more days of exercise. The majority (77.73%) of the power users used an iPhone versus Android, and a lower percentage were Web users as compared with the occasional or basic users. A higher proportion also (14.00%) had at least one or more devices/apps linked to the app versus the occasional users (3.70%) or the basic users (7.82%). The power users also had more friends on the app; were part of a group; had been an administrator of a challenge; and had more customized goals, foods, recipes, and exercises than the other subgroups. They also had a higher percentage of app customization (eg, app reminders, setting up a picture).

With respect to the robustness of the exploratory analyses, the AUC obtained from data mining validation sample 1 was 0.8327 (95% CI, 0.8306-0.8348), indicating good accuracy. The AUC obtained from data mining validation sample 2 was 0.8339 (95% CI, 0.8318-0.8359); thus, indicating high reliability. The CART model using data mining validation sample 2 is shown in Figure 3.

The factors used to predict the initial splits were almost identical to the model obtained from the training sample. Varying the complexity parameter in data mining validation sample 2 further subdivided the weight loss subgroups, based on food calories logged and weigh-ins. The overall model, however, is comparable to the initial model that used the training sample.

Based on the results that characterized the subgroups identified in the CART analyses, customization of the app appeared to be important among those who were more successful at losing weight. The group with a higher proportion of weight loss (the power users) used more features of the app than the other 2 weight loss subgroups. To explore the extent to which customization led to higher weight loss success, we conducted a logistic regression analysis post hoc using the training sample, with weight loss as the outcome and customization as the predictor. Weight loss was treated the same way as aforementioned, a dichotomous variable representing 5% or more of user's starting weight, and customization was derived as an ordinal variable consisting of 5 values (0-4 or more) that represented the number of customization features a user had (ie, whether a user had friends; was part of a group; administered a challenge; had custom goals, exercises, foods, or recipes; used reminders; used email reports; or had a picture). The odds of weight loss success progressively increased with more customization features compared with no customization features (1 customization feature: odds ratio, OR=5.27, 95% CI=5.11-5.44; 2 customization features: OR=12.39, 95% CI=11.99-12.81; 3 customization features: OR=22.42, 95% CI=21.56-23.31; 4 customization features: OR=48.30, 95% CI=46.23-50.46). Similar results were obtained with data mining validation sample 1.

Figure 3. Classification and regression tree for identifying successful weight loss subgroups with data mining validation sample 2 (n=323,975), varying the complexity parameter, minimum node split, and terminal node. Note: Factors for initial splits are similar to Figure 2. Subgroups from similar splits are bolded.
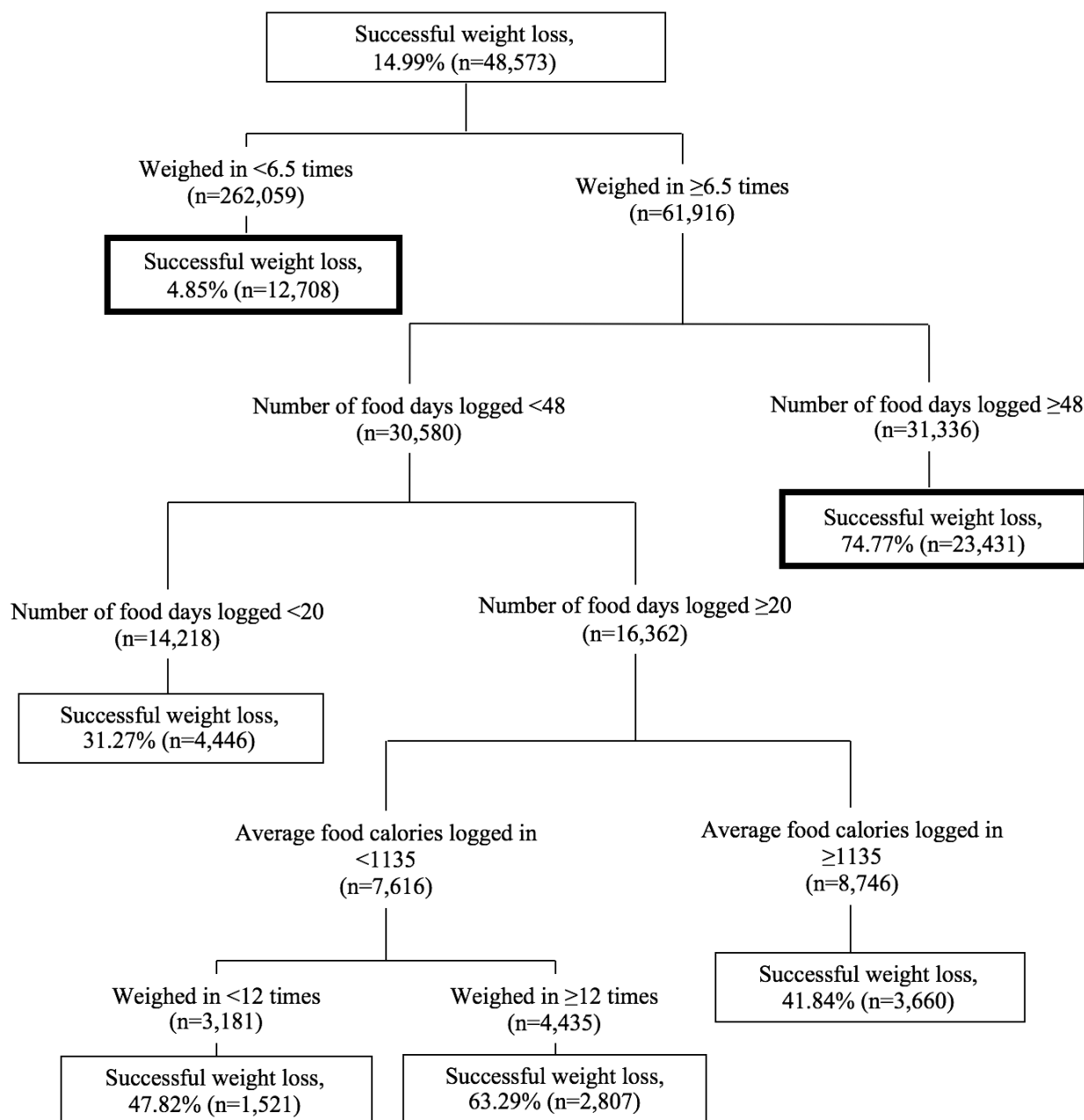
Discussion

Commercial weight loss apps can reach large segments of society, and data from these apps can provide possible clues to subgroups that are more or less successful at losing weight. However, these data can be messy and few researchers have attempted to systematically detect the signal from the noise with this type of data, using exploratory data mining methods. In addition to providing a model for exploring large quantities of commercially generated mobile health data, this study used analytic techniques to systematically examine the robustness and reliability of results obtained from exploratory analyses.

Results indicated key behavioral factors (eg, the number of times a user weighs in and the number of food days a user logs on the app) classified subgroups with varying proportions of weight loss success. On further exploration of characteristics of weight loss, users who were more successful at weight loss logged in about 8 times more days of exercise than the other subgroups. These findings are consistent with the literature demonstrating frequent self-monitoring, such as weighing in and logging in food and exercise, is associated with greater weight loss and decreased risk of weight regain [30-34].

Unexpectedly, this study found that the most successful weight loss subgroup (the power users) had a significantly higher number of iPhone users, compared with Android or Web users. Whether this is due to differences in iPhone versus other users or differences in the user experience of the app is unclear. Moreover, having friends on the app appears to be an important characteristic of weight loss, accounting for about 24% of the variance between subgroups. The power users had about 25%

more friends on the app than the occasional users. Studies have shown that social networks have become commonplace for individuals wanting to share information and seeking emotional support for issues regarding weight loss [35,36], and this is highly correlated with weight loss [37-40].

This study further suggests that greater customization of the app is associated with more likelihood of successful weight loss. Thus, although key behavioral factors are important in identifying more versus less successful weight loss subgroups, how users interact with the app may also be important. It may be possible that individuals who customize their app tend to be more engaged with their app, and those who are more engaged are more likely to be more motivated. This hypothesis warrants further investigation.

Limitations

There were a number of limitations associated with this study. First, the sample may not be representative of a national population. To examine this, we compared our entire app sample with a nationally representative sample (ie, 2008-2014 National Health Information Survey, NHIS, data) to examine differences. When we restricted both samples to include those aged only 18-70 years old (the app: n=10,444,981; NHIS: n=186,134 with replicate weights), we found that the app sample had a higher percentage of women (75.40%) than the NHIS sample (50.96%). The app sample was slightly younger (35.5 years) than the NHIS sample (42.6 years). When we applied both age and weight exclusion criteria to include only overweight and obese adults, these differences persisted, although the average BMIs were comparable between the 2 samples.

Second, the weight data were self-reported which may lead to inaccurate data. We examined the BMI values in the app sample with the NHIS sample where BMI is also calculated using self-reported data. The NHIS sample had a lower average BMI, 27.8, compared with the app sample where the average starting BMI was 30.4. When we examined only overweight and obese

adults, the app sample had only slightly higher starting BMI values than the NHIS sample (32.8 vs 31.0). Still, whether the results from this study can generalize to overweight and obese individuals more broadly is unknown.

Third, the data we analyzed were metadata and summary data. Therefore, we could only assess changes in weight at a general level, but not more specific longitudinal patterns. Thus, we could not assess more time-intensive longitudinal patterns of weight loss.

Conclusions

This study provides an approach to apply scientific methods to large health datasets collected by commercial apps and other health behavior technologies. Using both exploratory data mining and validation methods with big data in rapid fashion can increase confidence in the results that are obtained. Researchers should look to optimize scientific rigor, especially when trying to detect signal from noise in messy datasets.

In addition, the identification of particular subgroups that are successful at weight loss may help to inform researchers and practitioners involved in designing interventions with mobile technologies and smartphone apps. For example, weight loss interventions that use mobile technologies might aim to design interventions that emphasize behavioral factors and encourage individuals to customize their app experience. Furthermore, this study used data mining techniques that aid in hypothesis generation. Future studies should test the mechanisms underlying the behavior change, in this case, weight loss.

As more and more health app data become available, methods to analyze such big data will be crucial. Indeed, the era of big data offers new opportunities to better understand health behavior and behavior change, as well as potentially advance health behavior theories that help to explain mechanisms of behavior change. Our study provides an example for researchers to take full advantage of such opportunities.

Conflicts of Interest

None declared.

References

1. Anderson M. Pew Research Center. 2015. Technology Device Ownership: 2015 URL: http://www.pewinternet.org/2015/10/29/technology-device-ownership-2015/ [accessed 2016-04-03] [WebCite Cache ID 6gVHeTuP7]
2. Fox S, Duggan M. Pew Research Center. Mobile Health 2012 URL: http://www.pewinternet.org/2012/11/08/mobile-health-2012/ [accessed 2015-12-21] [WebCite Cache ID 6dwglz4aG]
3. Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of childhood and adult obesity in the United States, 2011-2012. JAMA 2014 Feb 26;311(8):806-814. [doi: 10.1001/jama.2014.732] [Medline: 24570244]
4. Carter MC, Burley VJ, Nykjaer C, Cade JE. Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. J Med Internet Res 2013;15(4):e32 [FREE Full text] [doi: 10.2196/jmir.2283] [Medline: 23587561]
5. Stephens J, Allen J. Mobile phone interventions to increase physical activity and reduce weight: a systematic review. J Cardiovasc Nurs 2013;28(4):320-329 [FREE Full text] [doi: 10.1097/JCN.0b013e318250a3e7] [Medline: 22635061]

XSL•FO

RenderX

6. Stephens J, Allen JK, Dennison Himmelfarb Cheryl R. "Smart" coaching to promote physical activity, diet change, and cardiovascular health. J Cardiovasc Nurs 2011;26(4):282-284 [FREE Full text] [doi: 10.1097/JCN.0b013e31821ddd76] [Medline: 21666418]

7. Pagoto S, Schneider K, Jojic M, DeBiasse M, Mann D. Evidence-based strategies in weight-loss mobile apps. Am J Prev Med 2013 Nov;45(5):576-582. [doi: 10.1016/j.amepre.2013.04.025] [Medline: 24139770]

8. Hammond RA, Levine R. The economic impact of obesity in the United States. Diabetes Metab Syndr Obes 2010;3:285-295 [FREE Full text] [doi: 10.2147/DMSOTT.S7384] [Medline: 21437097]

9. Wang YC, McPherson K, Marsh T, Gortmaker SL, Brown M. Health and economic burden of the projected obesity trends in the USA and the UK. The Lancet 2011 Aug;378(9793):815-825. [doi: 10.1016/S0140-6736(11)60814-3]

10. Withrow D, Alter D. The economic burden of obesity worldwide: a systematic review of the direct costs of obesity. Obes Rev 2011 Feb;12(2):131-141. [doi: 10.1111/j.1467-789X.2009.00712.x] [Medline: 20122135]

11. Behrens J. Principles and procedures of exploratory data analysis. Psychological Methods 1997;2(2):131-160. [doi: 10.1037/1082-989X.2.2.131]

12. Greenhouse J. Exploratory statistical methods, with applications to psychiatric research. Psychoneuroendocrinology 1992 Oct;17(5):423-441. [doi: 10.1016/0306-4530(92)90001-N]

13. Tukey J. Exploratory Data Analysis. Reading, MA: Addison-Wesley; 1977.

14. Raubertas RF, Rodewald LE, Humiston SG, Szilagyi PG. ROC curves for classification trees. Med Decis Making 1994;14(2):169-174. [Medline: 8028469]

15. Bradley A. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognition 1997 Jul;30(7):1145-1159. [doi: 10.1016/S0031-3203(96)00142-2]

16. Centers for Disease Control and Prevention. Losing Weight URL: http://www.cdc.gov/healthyweight/losing_weight/ [accessed 2015-12-21] [WebCite Cache ID 6dwgtMghx]

17. Mertens I, Van Gaal L F. Overweight, obesity, and blood pressure: the effects of modest weight reduction. Obes Res 2000 May;8(3):270-278 [FREE Full text] [doi: 10.1038/oby.2000.32] [Medline: 10832771]

18. Look AHEAD Research Group, Wing RR. Long-term effects of a lifestyle intervention on weight and cardiovascular risk factors in individuals with type 2 diabetes mellitus: four-year results of the Look AHEAD trial. Arch Intern Med 2010 Sep 27;170(17):1566-1575 [FREE Full text] [doi: 10.1001/archinternmed.2010.334] [Medline: 20876408]

19. Lemon S, Roy J, Clark M, Friedmann P, Rakowski W. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. Ann Behav Med 2003 Dec;26(3):172-181. [Medline: 14644693]

20. McArdle J. Exploratory data mining using CART in the behavioral sciences. In: Cooper H, Camic P, Long D, Panter A, Rindskopf D, Sher K, editors. APA Handbook of Research Methods in Psychology. Washington, DC: American Psychological Association; 2012:405-421.

21. Atienza AA, Yaroch AL, Mãsse LC, Moser RP, Hesse BW, King AC. Identifying sedentary subgroups: the National Cancer Institute's Health Information National Trends Survey. Am J Prev Med 2006 Nov;31(5):383-390 [FREE Full text] [doi: 10.1016/j.amepre.2006.07.024] [Medline: 17046409]

22. Dunton GF, Atienza AA, Tscherne J, Rodriguez D. Identifying combinations of risk and protective factors predicting physical activity change in high school students. Pediatr Exerc Sci 2011 Feb;23(1):106-121. [Medline: 21467595]

23. King A, Goldberg J, Salmon J, Owen N, Dunstan D, Weber D, et al. Identifying subgroups of U.S. adults at risk for prolonged television viewing to inform program development. Am J Prev Med 2010 Jan;38(1):17-26. [doi: 10.1016/j.amepre.2009.08.032] [Medline: 20117553]

24. Loh W. Classification and regression trees. WIREs Data Mining Knowl Discov 2011 Jan 06;1(1):14-23. [doi: 10.1002/widm.8]

25. Breiman L. Classification and regression trees. New York: Chapman & Hall; 1984.

26. Kraemer H. Evaluating medical tests: objective and quantitative guidelines. Newbury Park, CA: Sage Publications; 1992.

27. Lewis R. An Introduction to Classification and Regression Tree (CART) Analysis. 2000 Presented at: Annual Meeting of the Society for Academic Emergency Medicine; 2000; San Francisco, CA URL: http://www.researchgate.net/publication/240719582 [WebCite Cache ID 6i4Dbhm5H]

28. Therneau T, Atkinson E. Mayo Foundation. 2015 Jun. An Introduction to Recursive Partitioning Using the RPART Routines URL: https://cran.r-project.org/web/packages/rpart/vignettes/longintro.pdf [WebCite Cache ID 6dwgz2E6v]

29. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F. R-project. Package 'pROC' URL: https://cran.r-project.org/web/packages/pROC/pROC.pdf [accessed 2015-12-21] [WebCite Cache ID 6dwh187PZ]

30. Wharton C, Johnston C, Cunningham B, Sterner D. Dietary self-monitoring, but not dietary quality, improves with use of smartphone app technology in an 8-week weight loss trial. J Nutr Educ Behav 2014;46(5):440-444. [doi: 10.1016/j.jneb.2014.04.291] [Medline: 25220777]

31. Burke L, Conroy M, Sereika S, Elci O, Styn M, Acharya S, et al. The effect of electronic self-monitoring on weight loss and dietary intake: a randomized behavioral weight loss trial. Obesity (Silver Spring) 2011 Feb;19(2):338-344 [FREE Full text] [doi: 10.1038/oby.2010.208] [Medline: 20847736]

32. Burke L, Wang J, Sevick M. Self-monitoring in weight loss: a systematic review of the literature. J Am Diet Assoc 2011 Jan;111(1):92-102 [FREE Full text] [doi: 10.1016/j.jada.2010.10.008] [Medline: 21185970]

XSL•FO
RenderX

33. Linde J, Jeffery R, French S, Pronk N, Boyle R. Self-weighing in weight gain prevention and weight loss trials. Ann Behav Med 2005 Dec;30(3):210-216. [doi: 10.1207/s15324796abm3003_5] [Medline: 16336072]

34. Wing RR, Tate DF, Gorin AA, Raynor HA, Fava JL. A self-regulation program for maintenance of weight loss. N Engl J Med 2006 Oct 12;355(15):1563-1571. [doi: 10.1056/NEJMoa061883] [Medline: 17035649]

35. Ballantine P, Stephenson R. Help me, I'm fat! Social support in online weight loss networks. J. Consumer Behav 2011 Dec 23;10(6):332-337. [doi: 10.1002/cb.374]

36. Li V, McDonald D, Eikey E, Sweeney J, Escajeda J, Dubey G. Losing It Online: Characterizing Participation in an Online Weight Loss Community. USA; 2014 Presented at: GROUP '14 Proceedings of the 18th International ACM Conference on Supporting Group Work; 2014; Sanibel, FL p. 35-45. [doi: 10.1145/2660398.2660416]

37. 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 [FREE Full text] [doi: 10.1016/j.amjmed.2007.04.017] [Medline: 17602934]

38. Hwang KO, Ning J, Trickey AW, Sciamanna CN. Website usage and weight loss in a free commercial online weight loss program: retrospective cohort study. J Med Internet Res 2013;15(1):e11 [FREE Full text] [doi: 10.2196/jmir.2195] [Medline: 23322819]

39. 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]

40. Weinstein PK. A review of weight loss programs delivered via the Internet. J Cardiovasc Nurs 2006;21(4):251-8; quiz 259. [Medline: 16823276]

Abbreviations

BMI: body mass index
CART: Classification and regression tree
NHIS: National Health Information Survey

Original Paper

The Influence of Wireless Self-Monitoring Program on the Relationship Between Patient Activation and Health Behaviors, Medication Adherence, and Blood Pressure Levels in Hypertensive Patients: A Substudy of a Randomized Controlled Trial

Ju Young Kim[1,2], M.D., Ph.D; Nathan. E Wineinger[1], Ph. D; Steven. R Steinhubl[1], M.D.

[1]Digital medicine, Scripps Translational Science Institute, La Jolla, CA, United States

[2]Department of Family Medicine, Seoul National University Bundang Hospital, Seongnam, Republic Of Korea

Corresponding Author:
Steven. R Steinhubl, M.D.
Digital medicine
Scripps Translational Science Institute
3344 North Torrey Pines Court, Suite 300
La Jolla, CA, 92037
United States
Phone: 1 858 554 5708
Fax: 1 858 546 9284
Email: steinhub@scripps.edu

Abstract

Background: Active engagement in the management of hypertension is important in improving self-management behaviors and clinical outcomes. Mobile phone technology using wireless monitoring tools are now widely available to help individuals monitor their blood pressure, but little is known about the conditions under which such technology can effect positive behavior changes or clinical outcomes.

Objective: To study the influence of wireless self-monitoring program and patient activation measures on health behaviors, medication adherence, and blood pressure levels as well as control of blood pressure in hypertensive patients.

Methods: We examined a subset of 95 hypertensive participants from a 6-month randomized controlled trial designed to determine the utility of a wireless self-monitoring program (n=52 monitoring program, n=43 control), which consisted of a blood pressure monitoring device connected with a mobile phone, reminders for self-monitoring, a Web-based disease management program, and a mobile app for monitoring and education, compared with the control group receiving a standard disease management program. Study participants provided measures of patient activation, health behaviors including smoking, drinking, and exercise, medication adherence, and blood pressure levels. We assessed the influence of wireless self-monitoring as a moderator of the relationship between patient activation and health behaviors, medication adherence, and control of blood pressure.

Results: Improvements in patient activation were associated with improvements in cigarette smoking (beta=−0.46, P<.001) and blood pressure control (beta=0.04, P=.02). This relationship was further strengthened in reducing cigarettes (beta=−0.60, P<.001), alcohol drinking (beta=−0.26, P=.01), and systolic (beta=−0.27, P=.02) and diastolic blood pressure (beta=−0.34, P=.007) at 6 months among individuals participating in the wireless self-monitoring program. No differences were observed with respect to medication adherence.

Conclusions: Participation in a wireless self-monitoring program provides individuals motivated to improve their health management with an added benefit above and beyond that of motivation alone. Hypertensive individuals eager to change health behaviors are excellent candidates for mobile health self-monitoring..

Trial Registration: ClinicalTrials.gov NCT01975428, https://clinicaltrials.gov/ct2/show/NCT01975428 (Archived by WebCite at http://www.webcitation.org/6iSO5OgOG)

XSL•FO
RenderX

Introduction

Hypertension is the single greatest attributable risk factor for death and disease burden worldwide from noncommunicable disease [1]. Management of blood pressure has been shown to reduce the risk of stroke, heart attack, heart failure, and cardiovascular death [2,3]. However, less than half of young adults and two-thirds of older adults with treatment-eligible hypertension meet blood pressure goals [4].

The success of long-term effective control of blood pressure largely depends on patient self-management, where the individual plays a central role in their decisions and behaviors regarding management of this chronic condition. Self-management skills in chronic conditions include medication adherence, dietary compliance, self-monitoring of key parameters such as blood pressure, and coping skills to better manage anger or frustration that comes from living with disease [5,6]. A variety of tools to improve self-management have been evaluated as strategies for improving the treatment of a number of chronic conditions, most by emphasizing patient education and by creating a more active role of the patient in their health [7]. A critical part in self-management is active, engaged patients.

The Patient Activation Measure (PAM) is one of the tools for measuring an individual's skills, confidence, and knowledge in managing his or her own health [8,9]. PAM scores can be categorized into 4 levels: (1) the patient as a passive recipient of health care, (2) interested but lacking knowledge and confidence to act, (3) taking actions but lacking confidence and skills to support new behaviors, and (4) adoption of new behaviors but may not necessarily sustain them under crisis or stress.

Higher PAM scores have shown strong relationships with better performance of self-management behaviors, higher medication adherence, and higher quality-of-life score in patients with chronic condition [10]; clinical outcomes such as hemoglobin A_{1c} levels in diabetic patients [11]; and lower health care utilization, emergency room visits, and admissions [12]. PAM score has also been associated with changes in health behaviors, safety, cancer risk, stress, and mental health [13].

With the development of mobile medical technologies, tools to enhance self-management that combine home monitoring with near real-time bidirectional communication and feedback are becoming more widely available through mobile phones and wireless hubs [14]. The greatest strength of mobile health (mHealth) for patients may be in monitoring their disease [15]. Mobile health devices that measure physiological metrics allow individuals to track their health condition, diagnose health problems, or identify early warning signals without the direct need of a health care provider. A recent randomized clinical trial that used home blood pressure telemonitoring combined with automated self-care support using mobile phone found that mean daytime ambulatory systolic blood pressure decreased

significantly by more than 9 mm Hg in the intervention group [16]. Several clinical trials have also proven that self-monitoring of blood pressure, when coupled with health care provider feedback, can result in improved blood pressure control [17,18].

However, the mechanism of improved blood pressure control in individuals utilizing blood pressure self-monitoring is still not well understood. According to a meta-analysis of blood pressure self-monitoring on medication adherence and lifestyle factors from 28 trials with 7021 participants, monitoring was shown to have a small but significant effect on medication adherence but an inconclusive result related to diet and physical activities [19]. One other study found that using a monitoring system over 3 months improved PAM scores, self-care, and quality of life in elderly patients with heart failure [20]. However, another study found no relationship between PAM score and blood pressure in patients with uncontrolled hypertension receiving a comprehensive home care and monitoring system [21].

We originally performed the prospective, randomized controlled trial of the effect of wireless self-monitoring program on health care utilization in a heterogeneous group of patients with hypertension, type 1 or type 2 diabetes mellitus, or cardiac arrhythmia [22]. This study sought to determine whether wireless self-monitoring and PAM could be an effective way for improving health outcomes in hypertensive patients. Importantly, if wireless blood pressure self-monitoring programs can improve patient activation and downstream health behaviors, medication adherence, and blood pressure control, then strategies for such programs can be optimized to influence patient activation.

However, to date there have been limited studies of PAM, its modifiability through mHealth blood pressure self-monitoring, and their effects on behavioral factors, medication adherence, and blood pressure levels in a hypertensive population. Our study aimed to evaluate the interaction between wireless self-monitoring program and PAM in positive behavior changes, medication adherence, and control of blood pressure among hypertensive patients.

Methods

Design and Setting

This study was a secondary analysis of a prospective, randomized controlled, 2-group, pre-post intervention trial on the 6-month effectiveness of mobile phone–based self-monitoring program of chronic conditions on health care utilization [22]. The trial took place between July 2013 and December 2014, was approved by the Scripps Institutional Review Board (approval no: IRB-12-6019), and registered (trial registration: ClinicalTrials.gov identifier NCT01975428)].

In brief, this study was a collaboration between Scripps Translational Science Institute, Scripps Health, HealthyCircles by Qualcomm Life, and HealthComp, the third-party administrator for Scripps Health responsible for processing all

XSL•FO

RenderX

health care insurance claims for employees and dependents. Study participants in the primary trial were drawn from a pool of eligible employees and dependents insured by Scripps Health in 2012 and identified to have a *Current Procedural Terminology* (CPT) code related to one or more of the following diagnoses: hypertension, diabetes, or cardiac arrhythmia according to a medical claims database. The 3998 eligible individuals were identified according to inclusion criteria of being able to attend visits at a Scripps facility, being able to access the Internet, participating in HealthComp disease management program, and being willing to use wireless devices, iPhone, and the HealthyCircles Platform. They were also identified as the highest 25% in terms of their 2012 medical billing data to account for the highest health care spending and resource utilization. Individuals were recruited through an emailed invitation letter, followed up with a telephone call. If individuals expressed interest in the study, they were provided with a link to a Web-based informed consent form through email. After consenting to the study, study participants completed a Web-based survey that included demographic, health, and technology-related questions. Afterward, participants were randomized to control or intervention and brought in for an enrollment visit with an unblinded research coordinator, which they repeated at the conclusion of their 6-month study enrollment at a Scripps facility. Participants were blind to their assigned group before enrollment. All participants were randomized (ratio 1:1) to either intervention group, receiving wireless monitoring program plus disease management, or control group, receiving standard disease management program.

Our study focused only on those enrolled in the original study with a diagnosis of hypertension, prescribed at least one or more antihypertensive medications, and completed their 6-month follow-up study visit. The flow of study participants and subset of participants in this study are presented in Figure 1.

Figure 1. Study enrollment flowchart.
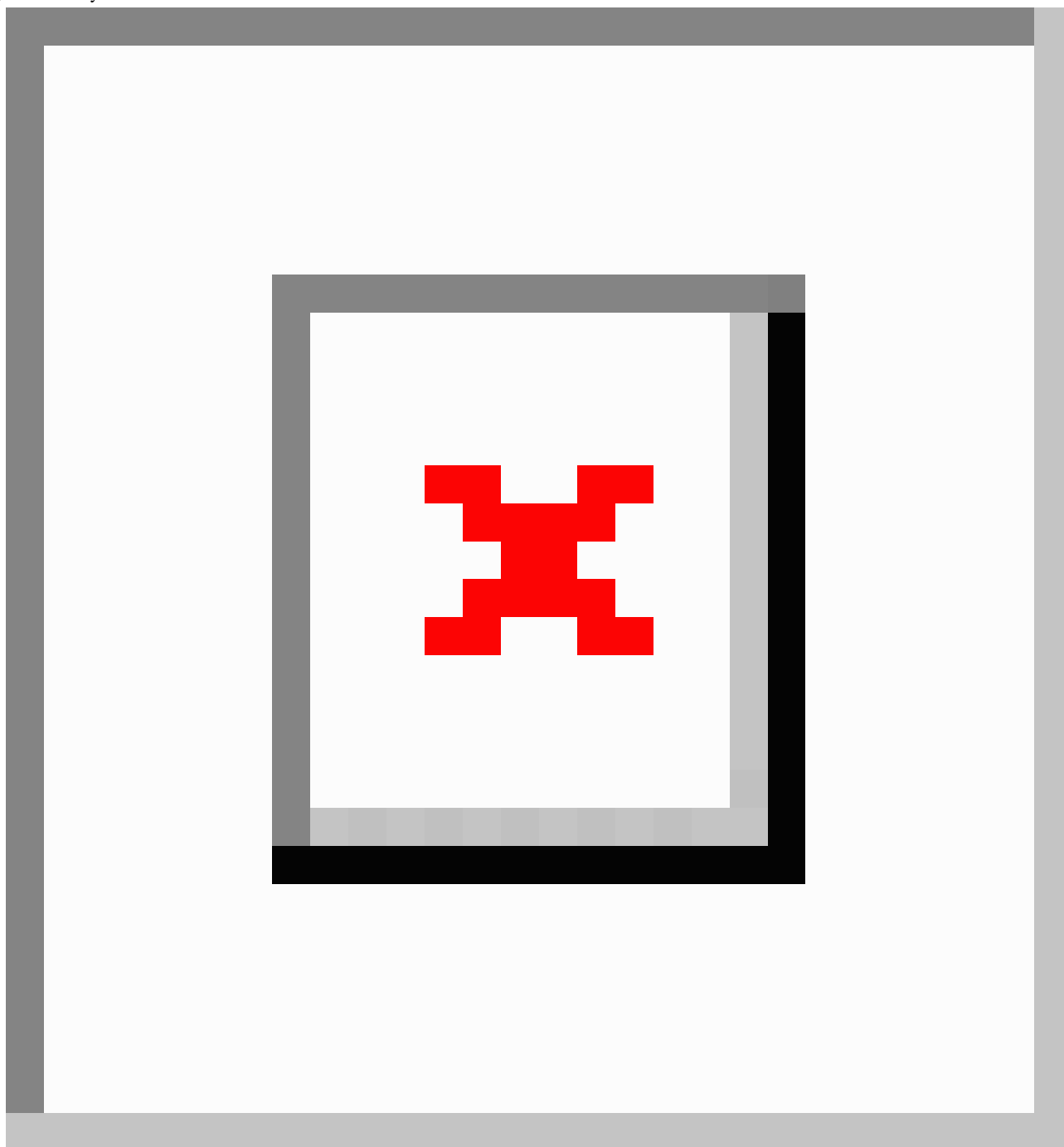
Wireless Self-Monitoring and Control Groups

During the study enrollment visit, all study participants completed the baseline survey if not already completed online and underwent height, weight, and blood pressure measurements, as well as a 12-lead electrocardiogram and fasting blood glucose test in the case of hypertensive individuals with a history of cardiac arrhythmia and diabetes, respectively.

Participants randomized to the monitoring group were provided with a study monitoring device according to their condition. Those enrolled with a diagnosis of hypertension received a Withings Blood Pressure Monitor. Participants in the monitoring group also received an iPhone with corresponding apps and were enrolled in the HealthyCircles Platform—an online disease management program featuring educational materials, consumer portals, and a dashboard to link with their families, caregivers, and health care professionals. HealthComp is the third-party administrator for Scripps Health, and HealthComp nursing staff had access to the HealthyCircles care management dashboard that showed the participant's device monitoring results and trends over time. Readings recorded on the devices were wirelessly uploaded to the HealthyCircles account and were accessible to the patient as well as the HealthComp nurses via an iPhone or computer. Also included in the management platform were reminders for monitoring, information about the disease condition, and general health behavior recommendations. Hypertensive participants in the monitoring group were trained on how to use their mobile phone, the HealthyCircles mobile

app, portal, and their device. They were encouraged to use the device 3 times a week, taking 2 measurements per day, with the first in the morning. If their monitoring frequency fell below 3 times per week for 2 consecutive weeks, participants received an email on their HealthyCircles Platform reminding them of the monitoring schedule. If participants were recommended more frequent monitoring by their physician than that asked through the study, they were encouraged to follow the physician's instructions. Figures 2 and 3 are screenshots of the HealthyCircles online portal and mobile app.

All participants, including participants randomized to control group, were enrolled in the HealthComp disease management program, and HealthComp nursing staff could reach out to all participants for the purpose of relaying medical education and wellness information with regard to disease prevention and chronic disease management.

Figure 2. Screenshot of HealthyCircles online portal in wireless monitoring program.

Figure 3. Screenshot of self-monitoring blood pressure data on mobile phone.
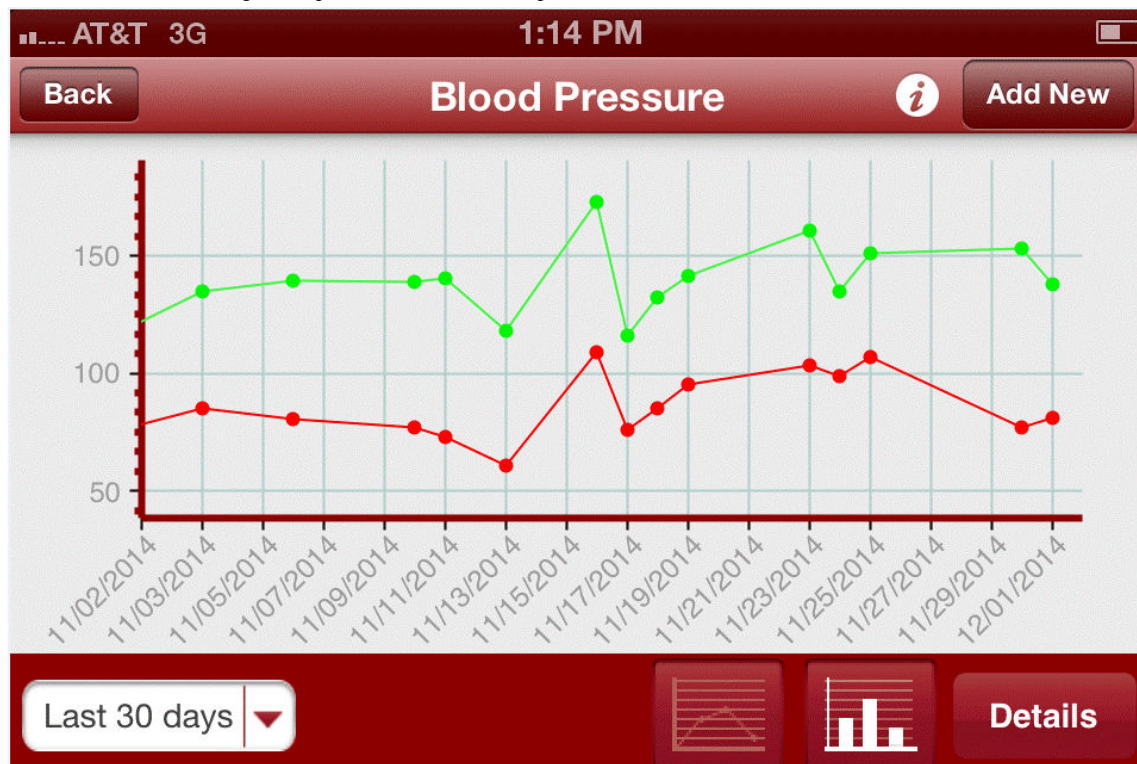
Outcome Measurements

Three health behaviors were assessed at both baseline and the 6-month follow-up Web-based survey: frequency of the use of alcohol, smoking, and exercise. Alcohol frequency was assessed by one survey item asking, "In the *past 6 months*, how many alcoholic drinks per week do you/did you have? (if you are a current or previous drinker)?" and the answer was coded as one of following categories: never, monthly or less, 2-4 times a month, 2-3 times a week, and 4 or more times a week. Smoking frequency was assessed by one survey item asking, "In the *past 6 months*, how many cigarettes/cigars per day do you/did you smoke? (if you are a current or previous user of tobacco)" and the answer was coded as one of the following categories: never, 10 or fewer, 11 to 20, 21 to 30, and 31 or more cigarettes. Finally, exercise frequency was assessed using the validated Godin Leisure-Time Exercise Questionnaire (GLTEQ) [23]. The GLTEQ consisted of average frequency of mild (minimal effort), moderate (not exhausting), and strenuous (producing rapid heartbeats) activities undertaken for more than 15 minutes during a typical week. Medication adherence was likewise assessed using a validated 8-item self-reported questionnaire [24]. Systolic and diastolic blood pressure outcomes were based on those measured by a research nurse using the Withings Blood Pressure Monitor at the baseline and study completion visits. Finally, individuals were classified as either having or not having adequate blood pressure control based on recommendations from the Eighth Joint National Committee (JNC 8) [25].

Determinants

The primary independent variables of interest were (1) assignment to the wireless self-monitoring or control group and (2) the relative change in the PAM score from baseline to the end of the study. Our primary hypothesis examined the moderator effect of this monitoring group assignment on the relative change in PAM on the outcome measures. That is, we jointly modeled the main effects of wireless self-monitoring and relative change in PAM, and the interaction between these variables. There was a problem regarding the database where a subset of Withings measurements could not be accurately determined, so self-monitoring was simply treated as a dichotomous variable.

Meanwhile, PAM was derived from an interval level, unidimensional, Guttman-like measure questionnaire consisting of 13 items. Each item has 4 possible responses ranging from "disagree strongly" (coded as 1) to "agree strongly" (coded as 4). An overall PAM score can be calculated by summing the coded responses. Overall PAM scores are often used to place individuals into 1 of 4 levels based on thresholds. However, as we were interested in changes in patient activation from baseline, we retained the overall scores and calculated the relative change in PAM as the difference in activation score between baseline and follow-up divided by baseline activation score.

Potential Confounders

To account for potential confounding factors, several sociodemographic factors were modeled as potential covariates: age, sex, education level and income level, baseline blood pressure, the number and types of medications (ie, lipid-lowering drug, glucose-lowering drug, insulin, use of antidepressant medication, and anxiolytic medication), and the frequency with which an individual visited a physician in the 6 months before starting the trial.

Statistical Analysis

Because the original trial aimed to evaluate health care resource utilization as measured by health insurance claims and visits to the hospital including office visits, emergency room visits, inpatient stays, and all visits during the study periods, sample size was determined to be powered (a priori) to detect a 1–office visit difference between the control and the monitoring group (assuming a standard deviation of 2 office visits). Univariate analyses were performed on each of the variables assessed. Multivariable regression was performed in a stepwise approach to determine the effects of self-monitoring, relative changes in PAM, and their interaction on behavior changes, medication adherence, blood pressure, and blood pressure control. In total, 3 models were constructed: (1) covariates only, (2) covariates and the main effects of self-monitoring and relative changes in PAM, and (3) covariates, main effects, and their interaction. Covariates were selected based on marginal associations ($P<.10$) with the outcomes. All data analyses were performed using R software, version 3.2.0 [26].

Results

Our analysis focused on those study participants with a hypertension history who were prescribed antihypertensive medication according to claims data, completed both surveys, and provided blood pressure measurements at both study visits (n=95). This includes 52 participants assigned to the monitoring group and 43 participants assigned to the control group.

Table 1 lists the baseline characteristics of study participants. The mean age of participants was 57.6 years, with most participants being female (65/95, 68%), Caucasian (76/95, 80%), with most completing college (78/95, 82%), and generally in the middle-income bracket. The majority of participants were also nonsmokers (68/95, 72%) with about half drinking less than once per week (52/95, 55%) and being an active exerciser by GLTEQ category (41/95, 43%). In terms of clinical factors, in addition to a recent history of hypertension, around one-third of the study participants had a recent history of diabetes and/or cardiac arrhythmia. The mean number of antihypertensive medications per individual at baseline was 2.0, with 26% (25/95) taking an oral glucose-lowering medication, 54% (51/95) took lipid-lowering medication, and 34% (32/95) took antidepressant medication. Consistent with the randomization, there were no differences in baseline characteristics between the control and self-monitoring groups that could be explained outside of chance.

Summary statistics of independent and dependent variables of interest are presented in Table 2. The average baseline PAM score across all participants was 78.0 and was not significantly different from the end of study scores (mean 76.0, *P*=.34). Meanwhile, the average number of cigarettes smoked per day

across self-identified smokers significantly decreased from 16.7 to 1.9, and the percentage of nonsmokers increased from 72% to 96%. However, the average frequency of alcohol use (baseline: 5.5 drinks per month, end of study: 5.7; *P*=.94) and exercise group by GLTEQ categories (baseline: 36.8, end of study: 40.5; *P*=.17) did not change over the course of the study. Baseline Morisky Medication Adherence Scale (MAS) scores also did not change over the study period (baseline: 6.5, end of study: 6.6; *P*=.46). However, the average systolic and diastolic blood pressure levels decreased across all study participants over the study period (baseline: 140.6/89.4 mm Hg; end of study: 136.5/83.9 mm Hg). Furthermore, according to JNC 8 guidelines, the frequency of achieved blood pressure control increased from 45% to 59%.

Results of the multivariable models examining the relationship with PAM, wireless self-monitoring, and their interaction with the outcome variables of interest are presented in Table 3. The interaction between wireless self-monitoring and positive changes of PAM was a significant contributor to cigarette smoking (beta = −0.60, *P*<.001), alcohol drinking (beta = −0.26, *P*=.01), systolic blood pressure (beta = −0.27, *P*=.02), and diastolic blood pressure (beta = −0.34, *P*=.007) at 6 months. In all these cases, the wireless self-monitoring group had greater decreases in these outcome variables than the control group. This interaction was not a significant predictor of change in exercise per week, Morisky MAS, or achieved blood pressure control—although the main effects of both self-monitoring and changes in PAM were conditionally associated with achieved blood pressure control.

In order to further dissect these interactions, we evaluated the relationship between relative changes in PAM and number of cigarettes smoked per day, frequency of alcohol drinking, and systolic and diastolic blood pressure in the self-monitoring group and control group separately. We observed that increases in PAM scores were significantly associated with a decrease in smoking (beta = −0.63, *P*<.001), frequency of alcohol drinking (beta = −0.22, *P*=.04), systolic blood pressure (beta = −0.47, *P*<.001), and diastolic blood pressure (beta = −0.42, *P*<.001) in the monitoring group. However, no association was observed in the control group: smoking (beta = −0.11, *P*=.55), frequency of alcohol drinking (beta = 0.13, *P*=.48), systolic blood pressure (beta = 0.10, *P*=.46), or diastolic blood pressure (beta = −0.01, *P*=.93).

We also evaluated the temporal trends of frequency in self-monitoring of blood pressure as well as mean systolic and diastolic blood pressure levels in a subset of individuals (33/52) whose self-monitoring data were available during the total study period (Multimedia Appendix 1).

After 12 weeks, about 70% (23/33) had monitored their blood pressure regularly at a recommended frequency (median 6 per week, interquartile range [IQR] 5-9), and at the end of the study period, 55% (18/33) had still participated in the self-monitoring program with a median frequency of 6 per week (IQR 5-7; Multimedia Appendix 1). Baseline systolic and diastolic blood pressure levels in self-monitoring database were 133.6 mm Hg and 82.9 mm Hg, respectively. At the end of the study, systolic blood pressure was 129.7 mm Hg and diastolic was 78.4 mm Hg, showing decreased pattern over time but not a clinically meaningful decrease in terms of controlled hypertension.

Table 1. Baseline characteristics of hypertensive patients (n=95).

Characteristics		Monitoring (n=52)	Control (n=43)	Total	P value
Sociodemographic factors					
Age in years, mean (SD)		57.5 (8.6)	57.7 (8.7)	57.6 (8.6)	.89
Sex, n (%)					
	Female	38 (73)	27 (63)	65 (68)	.28
Ethnicity, n (%)					.07
	Caucasian	43 (83)	33 (77)	76 (80)	
	African American	4 (8)	2 (5)	6 (6)	
	Hispanic	4 (8)	1 (2)	5 (5)	
	Asian	1 (2)	4 (9)	5 (5)	
Married, n (%)		29 (56)	30 (70)	59 (62)	.40
Level of education, n (%)					
	≤12 years	6 (12)	11(26)	17 (18)	.09
	Complete college	22 (42)	20 (47)	42 (44)	
	More than college	24 (46)	12 (28)	36 (38)	
Income[a], n (%)					
	<50K	5 (10)	5 (12)	10 (11)	.81
	50-149K	35 (67)	32 (74)	67 (70)	
	150-249K	10 (19)	1 (2)	11 (12)	
	≥250K	5 (12)	7 (8)		
		2 (4)			
Alcohol, n (%)					
	None	7 (13)	11(26)	18 (19)	.22
	Less than 1/week	28 (54)	24 (56)	52 (55)	
Smoking, n (%)					
	None	32 (62)	36 (84)	68 (72)	.07
Exercise, n (%)					
	Active	23 (44)	18 (42)	41 (43)	.82
Clinical factors					
Number of antihypertensive medication		1.9 (1.0)	2.1 (1.0)	2.0 (1.0)	.38
Self-reported frequency of physician clinic visits in the past 6 months[a]		3.0 (4.4)	2.6 (1.7)	2.9(2.5)	.56
Comorbidity, n (%)					
	Type 1 Diabetes Mellitus	7 (13)	4 (9)	11 (12)	.53
	Type 2 Diabetes Mellitus	4 (8)	6 (14)	10 (11)	.32
	Arrhythmia	4 (8)	8 (19)	12 (13)	.11
Comedication, n (%)					
	Insulin	7 (13)	5 (12)	12 (13)	.79
	Diabetic medication	12 (23)	13 (30)	25 (26)	.43
	Lipid-lowering medication	25 (48)	26 (60)	51 (54)	.23
	Antidepressant	18 (35)	14 (33)	32 (34)	.83
	Anxiolytics	5 (10)	6 (14)	11 (12)	.51

[a]Variable with significant difference between monitoring group and control group.

Table 2. Patient activationmeasure, health behaviors, medication adherence, and blood pressure measures.

Characteristics	Baseline, mean (SD)	After 6 months, mean (SD)	P value
Predictors			
Patient Activation Measure			
Self-monitoring group	82.9 (13.8)	79.4 (21.9)	.19
Control group	72.0 (16.4)	72.1 (18.7)	.96
Outcome measurements			
Health behaviors			
Cigarettes per day among current smokers			
Self-monitoring group	16.5 (9.3)	2.6 (7.3)	<.001
Control group	17.1 (7.6)	0.3 (1.6)	.03
Frequency of drinking per month among drinkers			
Self-monitoring group	7.2 (8.0)	7.6 (7.9)	.66
Control group	6.2 (7.7)	5.8 (6.8)	.69
Exercise units per week			
Self-monitoring group	37.8 (26.1)	39.8 (23.8)	.58
Control group	35.6 (26.6)	41.3 (29.5)	.18
Morisky MAS [a]			
Self-monitoring group	6.6 (1.4)	6.7 (1.4)	.79
Control group	6.3 (1.4)	6.5 (1.5)	.47
Achieved blood pressure control, n (%)			
Self-monitoring group	29 (56)	34 (65)	.23
Control group	14 (33)	22 (51)	.001
Systolic blood pressure			
Self-monitoring group	136.1 (15.2)	133.4 (12.9)	.28
Control group	145.9 (19.5)	140.2 (18.4)	.06
Diastolic blood pressure			
Self-monitoring group	86.3 (12.8)	82.8 (11.2)	.06
Control group	93.1 (14.1)	85.3 (12.1)	.001

[a]Morisky MAS: Morisky Medication Adherence Scale 8 item.

Table 3. Multivariable regression model parameters estimating the effect of changes in PAM, wireless self-monitoring, and their interaction on the changes in health behaviors, medication adherence, and blood pressure at 6 months in all hypertensive patients (standardized regression coefficients and R^2 change with significance).

Model	Smoking (cigarettes/day)	Alcohol (frequency/month)	Exercise (unit/week)	Morisky MAS[a]	Systolic BP[b]	Diastolic BP	Achieved BP control
Model 1							
	R^2=0.15	R^2=0.60	R^2=0.39	R^2=0.47	R^2=0.29	R^2=0.20	R^2=0.36
Covariates							
Model 2							
	ΔR^2= 0.21	ΔR^2= 0.02	ΔR^2= 0.00	ΔR^2= 0.00	ΔR^2= 0.02	ΔR^2= 0.03	ΔR^2= 0.07
Δ PAM[c]	−0.46; P<.001	−0.05; P=.56	0.01; P=.89	0.02; P=.85	−0.12; P=.25	−0.25; P=.06	0.04; P=.02
Self-monitoring	0.04; P=.84	0.10; P=.19	−0.03; P=.78	−0.03; P=.60	−0.18; P=.07 −0.06; P=.50	0.6; P=.03	
Model 3							
	ΔR^2= 0.21	ΔR^2= 0.03	ΔR^2= 0.02	ΔR^2= 0.00	ΔR^2= 0.05	ΔR^2= 0.06	ΔR^2= 0.00
PAM*self-monitoring	−0.60; P<.001	−0.26; P=.01	0.15; P=.13	0.07; P=.48	−0.27; P=.02	−0.34; P=.007	0.01; P=.12

[a]Morisky MAS: Morisky Medication Adherence Scale 8 item.

[b]BP: blood pressure; PAM: Patient Activation Measure (13 items).

[c]Δ PAM: difference in PAM in 6 months/baseline PAM.

Discussion

Principal Findings

One goal of many chronic disease management programs is to adopt patient activation enhancing strategies in management of chronic disease. Such programs are generally designed to enhance problem solving, decision making, and taking actions [27]. These characteristics are common behavioral components of mobile self-monitoring and health intervention apps [28]. This study explored the moderating effects of wireless self-monitoring on relationships between changes of PAM score and changes of health behaviors, medication adherence, and blood pressure in individuals with hypertension over a 6-month period. The wireless self-monitoring program consisted of mobile phone–based self-monitoring with an online Web portal used for social support, education, reminders, and to track recordings, because this program was designed to leverage collaborations with device manufacturers, a connected health leader, health care provider, as well as employee wellness program.

We discovered that changes in PAM were associated decreases in the number of cigarettes smoked, alcohol consumed, and blood pressure in the self-monitoring group alone. This suggests wireless self-monitoring program may reinforce changes of PAM by enhancing the ability to make healthy decisions and to manage blood pressure.

To date there have been few studies that have explored the relationship between wireless self-monitoring and changes of PAM in chronic care. One study examining African Americans with uncontrolled hypertension showed that blood pressure

self-monitoring did not have any effect on changes of PAM [21]. Another study suggested that online social networks had favorable effects on changes of PAM in individuals with lower activation scores [29]. However, the study did not have the component of online self-monitoring in chronic disease population. Among 95 of our study's participants, 72 (76%) would be placed in the highest level of patient activation while 86 (91%) of participants were in the patient activation level 3 or higher. Thus, our study results support the finding that wireless self-monitoring can strengthen improved changes in PAM, even in the highest levels of patient activation [13].

We further discovered that changes in PAM were conditionally associated with changes in health behaviors and blood pressure control. These results support the finding that an increase in PAM score may improve health outcomes and potentially decrease long-term medical costs [30]. Importantly, this benefit appeared to be exclusive to individuals in the wireless self-monitoring group, suggesting that future strategies should strongly consider implementing mHealth in chronic disease management.

Despite this relationship between changes in PAM and blood pressure control, changes in PAM were not significantly associated with medication adherence. We observed a median Morisky MAS score of 7.0, which indicated relatively medium to high adherence. Our study participants consisted of Scripps Health hospital employees or their families. Thus, participants were likely to have higher health literacy or more availability of health personnel resources than the general population, which could possibly lead to better control of their condition. Furthermore, study participants generally had well-controlled hypertension, because the study participants took on average 2

antihypertensive medications and had a baseline blood pressure of 140.6/89.4 mm Hg, indicating generally mild to moderate hypertension. Finally, one previous study showed that uncontrolled depression was a barrier to antihypertensive medication adherence and was negatively associated with PAM [31]. Although one-third of our study's participants had been prescribed antidepressant medication, there is a high possibility that depression symptoms were also well controlled in our study. Further studies should examine these relationships in higher-risk populations.

The mechanism by which a wireless self-monitoring program potentiates the effect of changes in PAM on health behaviors and outcomes in hypertensive participants remains unknown. Although many studies have shown that patient activation level is associated with improved health behaviors and clinical outcomes [11,13,30,32], studies regarding effective intervention to enhance patient activation are limited. One prior study showed that remote monitoring with alert system increased activation levels in 21 patients with heart failure over 3 months [33].

In this study, the wireless self-monitoring program had integrative multiple components including online portal with social support from their family or friends, reminder for monitoring blood pressure, and education about the disease, as well as recommendation for positive behavior change for better outcomes. A study showed that Internet-based behavior intervention grounded in theory proved to be effective [34]. Among several proven theories, external supports such as personalized email or reminders to engage users in the intervention and providing peer support may improve adherence to specific intervention and outcomes in improving type 2 diabetes [35]. The availability of personal health-related data to individuals enrolled in the wireless self-monitoring group could have reinforced these changes. However, there is a great need for proven mHealth evidences in preventing cardiovascular disease [36].

In this study, Withings blood pressure monitoring device was used to track patient's blood pressure. Portable wireless device that enables monitoring of user's condition in clinical setting or off-site locations has its potential in improving health care with real-time monitoring and convenience. However, there are also potential risks if misused or misinterpreted. In 2015, the US Food and Drug Administration updated the 2013 document titled "Mobile Medical Applications: Guidance for Industry and Food and Drug Administration Staff" [37]. According to the document, if a blood pressure monitoring device was used for self-management of hypertension without providing specific treatment or treatment suggestion, it would not be regulated under the Federal Food, Drug, and Cosmetic Act [38]. However, if this device was intended for patient-specific analysis and providing patient-specific diagnosis or treatment recommendation, which seemed to be close to the next step of mHealth, it should be treated as a regulated medical device because it could pose risk to the general public. Moreover, reliability, validity, and preparation of regulatory pathway in mobile medical devices should be considered early during app design phase to ensure patient safety [39]. In addition, data security, communication between patients and health care providers, and usability should be discussed in a collaborative way between several shareholders of mobile medical apps.

Limitations

Although more than half the study participants enrolled in this trial were among the top quartile of eligible participants with respect to health care utilization, study participants exhibited relatively high levels of baseline medication adherence and PAM as described previously. The high baseline PAM levels among participants in our study prevented us from investigating the relationship between PAM, wireless self-monitoring, and study outcomes in individuals with low PAM scores.

In addition, the study participants were highly educated, were non-Hispanic white, and a female-dominant population compared with a prevalence study in the United States using National Health and Nutrition Examination Survey [40]. Baseline controlled hypertension group was 56% in self-monitoring group and 33% in control group, which could reduce the effect of wireless self-monitoring program. These factors limit generalizability of this study. Further research will be needed especially in uncontrolled hypertensive population.

Furthermore, a formal evaluation of the engagement in monitoring activities was not included, although we tried to look at the trends of monitoring frequency and mean systolic and diastolic blood pressure levels in a subset of participants. Over the 24 weeks of the study period, more than half of the participants were actively engaged in the self-monitoring program with recommended frequency, which was different in results of decreased adherence to protocol over time for a weight-monitoring app [41]. This might be partly due to the different perspective in weight monitoring and blood pressure monitoring, where blood pressure was considered a more significant burden to patient's health. In addition, automatic reminders and direct care from nursing staff of the disease management program could affect our usage over time.

Identifying the individual characteristics of mHealth self-monitoring programs that influence behavior will lead to improved systems. Finally, as study participants only participated for 6 months, further studies will be required to evaluate long-term benefit.

Conclusions

We discovered that a wireless self-monitoring program in hypertensive patients strengthened the relationship between positive changes in PAM and health behaviors and blood pressure control.

Adoption of wireless self-monitoring programs has the potential to enhance strategies for the management of hypertension. Further research should seek to identify individuals who may receive the most benefit from wireless self-monitoring programs.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Number of participants and mean systolic and diastolic blood pressure levels over time.

[PNG File, 20KB - jmir_v18i6e116_app1.png]

References

1. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, Des Jarlais Don C, Fowkes Francis Gerry R, Parry Charles D H, Stapelberg Nicolas J C, Murray Christopher J L, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2224-2260 [FREE Full text] [doi: 10.1016/S0140-6736(12)61766-8] [Medline: 23245609]

2. Blood Pressure Lowering Treatment Trialists' Collaboration, Sundström J, Arima H, Woodward M, Jackson R, Karmali K, et al. Blood pressure-lowering treatment based on cardiovascular risk: a meta-analysis of individual patient data. Lancet 2014 Aug 16;384(9943):591-598. [doi: 10.1016/S0140-6736(14)61212-5] [Medline: 25131978]

3. Sundström J, Arima H, Jackson R, Turnbull F, Rahimi K, Chalmers J, Blood Pressure Lowering Treatment Trialists' Collaboration. Effects of blood pressure reduction in mild hypertension: a systematic review and meta-analysis. Ann Intern Med 2015 Feb 3;162(3):184-191. [doi: 10.7326/M14-0773] [Medline: 25531552]

4. Navar-Boggan AM, Pencina MJ, Williams K, Sniderman AD, Peterson ED. Proportion of US adults potentially affected by the 2014 hypertension guideline. JAMA 2014 Apr 9;311(14):1424-1429. [doi: 10.1001/jama.2014.2531] [Medline: 24682242]

5. Lorig KR, Holman H. Self-management education: history, definition, outcomes, and mechanisms. Ann Behav Med 2003 Aug;26(1):1-7. [Medline: 12867348]

6. Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. JAMA 2002 Nov 20;288(19):2469-2475. [Medline: 12435261]

7. Grady PA, Gough LL. Self-management: a comprehensive approach to management of chronic conditions. Am J Public Health 2014 Aug;104(8):e25-e31. [doi: 10.2105/AJPH.2014.302041] [Medline: 24922170]

8. 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 [FREE Full text] [doi: 10.1111/j.1475-6773.2004.00269.x] [Medline: 15230939]

9. 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 [FREE Full text] [doi: 10.1111/j.1475-6773.2005.00438.x] [Medline: 16336556]

10. 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 Mar;30(1):21-29. [Medline: 17170635]

11. 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]

12. Begum N, Donald M, Ozolins IZ, Dower J. Hospital admissions, emergency department utilisation and patient activation for self-management among people with diabetes. Diabetes Res Clin Pract 2011 Aug;93(2):260-267. [doi: 10.1016/j.diabres.2011.05.031] [Medline: 21684030]

13. Harvey L, Fowles JB, Xi M, Terry P. When activation changes, what else changes? the relationship between change in patient activation measure (PAM) and employees' health status and health behaviors. Patient Educ Couns 2012 Aug;88(2):338-343. [doi: 10.1016/j.pec.2012.02.005] [Medline: 22459636]

14. Jones KR, Lekhak N, Kaewluang N. Using mobile phones and short message service to deliver self-management interventions for chronic conditions: a meta-review. Worldviews Evid Based Nurs 2014 Apr;11(2):81-88. [doi: 10.1111/wvn.12030] [Medline: 24597522]

15. Logan AG. Transforming hypertension management using mobile health technology for telemonitoring and self-care support. Can J Cardiol 2013 May;29(5):579-585. [doi: 10.1016/j.cjca.2013.02.024] [Medline: 23618506]

XSL•FO
RenderX

16.   Logan AG, Irvine MJ, McIsaac WJ, Tisler A, Rossos PG, Easty A, et al. Effect of home blood pressure telemonitoring with self-care support on uncontrolled systolic hypertension in diabetics. Hypertension 2012 Jul;60(1):51-57 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.111.188409] [Medline: 22615116]

17.   Hosseininasab M, Jahangard-Rafsanjani Z, Mohagheghi A, Sarayani A, Rashidian A, Javadi M, et al. Self-monitoring of blood pressure for improving adherence to antihypertensive medicines and blood pressure control: a randomized controlled trial. Am J Hypertens 2014 Nov;27(11):1339-1345 [FREE Full text] [doi: 10.1093/ajh/hpu062] [Medline: 24771706]

18.   Uhlig K, Patel K, Ip S, Kitsios GD, Balk EM. Self-measured blood pressure monitoring in the management of hypertension: a systematic review and meta-analysis. Ann Intern Med 2013 Aug 6;159(3):185-194. [doi: 10.7326/0003-4819-159-3-201308060-00008] [Medline: 23922064]

19.   Fletcher BR, Hartmann-Boyce J, Hinton L, McManus RJ. The Effect of Self-Monitoring of Blood Pressure on Medication Adherence and Lifestyle Factors: A Systematic Review and Meta-Analysis. Am J Hypertens 2015 Oct;28(10):1209-1221. [doi: 10.1093/ajh/hpv008] [Medline: 25725092]

20.   Evangelista LS, Lee J, Moore AA, Motie M, Ghasemzadeh H, Sarrafzadeh M, et al. Examining the effects of remote monitoring systems on activation, self-care, and quality of life in older patients with chronic heart failure. J Cardiovasc Nurs 2015;30(1):51-57 [FREE Full text] [doi: 10.1097/JCN.0000000000000110] [Medline: 24365871]

21.   Ryvicker M, Feldman PH, Chiu Y, Gerber LM. The role of patient activation in improving blood pressure outcomes in Black patients receiving home care. Med Care Res Rev 2013 Dec;70(6):636-652. [doi: 10.1177/1077558713495452] [Medline: 23864112]

22.   Bloss C, Wineinger N, Peters M, Boeldt D, Ariniello L, Kim J, et al. A prospective randomized trial examining health care utilization in individuals using multiple smartphone-enabled biosensors. PeerJ 2016;4:e1554 [FREE Full text] [doi: 10.7717/peerj.1554] [Medline: 26788432]

23.   Godin G, Shephard RJ. A simple method to assess exercise behavior in the community. Can J Appl Sport Sci 1985 Sep;10(3):141-146. [Medline: 4053261]

24.   Morisky DE, Ang A, Krousel-Wood M, Ward HJ. Predictive validity of a medication adherence measure in an outpatient setting. J Clin Hypertens (Greenwich) 2008 May;10(5):348-354 [FREE Full text] [Medline: 18453793]

25.   James PA, Oparil S, Carter BL, Cushman WC, Dennison-Himmelfarb C, Handler J, et al. 2014 evidence-based guideline for the management of high blood pressure in adults: report from the panel members appointed to the Eighth Joint National Committee (JNC 8). JAMA 2014 Feb 5;311(5):507-520. [doi: 10.1001/jama.2013.284427] [Medline: 24352797]

26.   R-project for Statistical computing database on the Internet. URL: https://www.r-project.org/ [WebCite Cache ID 6dfi0KNFo]

27.   Farrell K, Wicks MN, Martin JC. Chronic disease self-management improved with enhanced self-efficacy. Clin Nurs Res 2004 Nov;13(4):289-308. [doi: 10.1177/1054773804267878] [Medline: 15448281]

28.   Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR Mhealth Uhealth 2015;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]

29.   Magnezi R, Bergman YS, Grosberg D. Online activity and participation in treatment affects the perceived efficacy of social health networks among patients with chronic illness. J Med Internet Res 2014;16(1):e12 [FREE Full text] [doi: 10.2196/jmir.2630] [Medline: 24413148]

30.   Greene J, Hibbard JH, Sacks R, Overton V, Parrotta CD. When patient activation levels change, health outcomes and costs change, too. Health Aff (Millwood) 2015 Mar 1;34(3):431-437. [doi: 10.1377/hlthaff.2014.0452] [Medline: 25732493]

31.   Eze-Nliam CM, Thombs BD, Lima BB, Smith CG, Ziegelstein RC. The association of depression with adherence to antihypertensive medications: a systematic review. J Hypertens 2010 Sep;28(9):1785-1795 [FREE Full text] [doi: 10.1097/HJH.0b013e32833b4a6f] [Medline: 20531223]

32.   Mitchell SE, Gardiner PM, Sadikova E, Martin JM, Jack BW, Hibbard JH, et al. Patient activation and 30-day post-discharge hospital utilization. J Gen Intern Med 2014 Feb;29(2):349-355 [FREE Full text] [doi: 10.1007/s11606-013-2647-2] [Medline: 24091935]

33.   Shively MJ, Gardetto NJ, Kodiath MF, Kelly A, Smith TL, Stepnowsky C, et al. Effect of patient activation on self-management in patients with heart failure. J Cardiovasc Nurs 2013;28(1):20-34. [doi: 10.1097/JCN.0b013e318239f9f9] [Medline: 22343209]

34.   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]

35.   Cotter AP, Durant N, Agne AA, Cherrington AL. Internet interventions to support lifestyle modification for diabetes management: a systematic review of the evidence. J Diabetes Complications 2014;28(2):243-251 [FREE Full text] [doi: 10.1016/j.jdiacomp.2013.07.003] [Medline: 24332469]

36.   Burke LE, Ma J, Azar Kristen M J, Bennett GG, Peterson ED, Zheng Y, American Heart Association Publications Committee of the Council on EpidemiologyPrevention, Behavior Change Committee of the Council on Cardiometabolic Health, Council on CardiovascularStroke Nursing, Council on Functional GenomicsTranslational Biology, Council on Quality of CareOutcomes Research, Stroke Council. Current Science on Consumer Use of Mobile Health for Cardiovascular Disease

XSL·FO

RenderX

Prevention: A Scientific Statement From the American Heart Association. Circulation 2015 Sep 22;132(12):1157-1213. [doi: 10.1161/CIR.0000000000000232] [Medline: 26271892]

37. U.S. Department of Health and Human Services Food and Drug Administration, Center for Devices and Radiological Health, Center for Biologics Evaluation and Research. FDA. Mobile Medical Applications : Guidance for Industry and Food and Drug Administration Staff URL: http://www.fda.gov/downloads/MedicalDevices/.../UCM263366.pdf [accessed 2016-04-29] [WebCite Cache ID 6h92m0iDI]

38. U.S. Food and Drug Administration. Legcounsel. FEDERAL FOOD, DRUG, AND COSMETIC ACT URL: http://legcounsel. house.gov/Comps/FDA_CMD.pdf [accessed 2016-04-29] [WebCite Cache ID 6h93XXc2q]

39. Yetisen AK, Martinez-Hurtado JL, da Cruz Vasconcellos Fernando, Simsekler M C Emre, Akram MS, Lowe CR. The regulation of mobile medical applications. Lab Chip 2014 Mar 7;14(5):833-840. [doi: 10.1039/c3lc51235e] [Medline: 24425070]

40. Yoon Sung Sug Sarah, Carroll MD, Fryar CD. Hypertension Prevalence and Control Among Adults: United States, 2011-2014. NCHS Data Brief 2015 Nov(220):1-8 [FREE Full text] [Medline: 26633197]

41. Carter MC, Burley VJ, Nykjaer C, Cade JE. Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. J Med Internet Res 2013;15(4):e32 [FREE Full text] [doi: 10.2196/jmir.2283] [Medline: 23587561]

Abbreviations

GLTEQ: Godin Leisure-Time Exercise Questionnaire
IQR: interquartile range
JNC 8: Eighth Joint National Committee
mHealth: mobile health
Morisky MAS: Morisky Medication Adherence Scale 8 item
PAM: Patient Activation Measure

Original Paper

Feasibility and Limitations of Vaccine Two-Dimensional Barcoding Using Mobile Devices

Cameron Bell[1], BEng; Julien Guerinet[1], BEng; Katherine M Atkinson[1,2], BSc; Kumanan Wilson[1,3], MD, MSc, FRCP(C)

[1]Ottawa Hospital Research Institute, Clinical Epidemiology Program, Ottawa, ON, Canada

[2]Karolinska Institutet, Department of Public Health Sciences, Stockholm, Sweden

[3]University of Ottawa, Departments of Medicine, Epidemiology and Community Medicine, Ottawa, ON, Canada

Corresponding Author:
Kumanan Wilson, MD, MSc, FRCP(C)
Ottawa Hospital Research Institute
Clinical Epidemiology Program
1053 Carling Ave.
Ottawa, ON, K1Y 4E9
Canada
Phone: 1 613 798 5555 ext 17921
Fax: 1 613 761 5492
Email: kwilson@ohri.ca

Abstract

Background: Two-dimensional (2D) barcoding has the potential to enhance documentation of vaccine encounters at the point of care. However, this is currently limited to environments equipped with dedicated barcode scanners and compatible record systems. Mobile devices may present a cost-effective alternative to leverage 2D vaccine vial barcodes and improve vaccine product-specific information residing in digital health records.

Objective: Mobile devices have the potential to capture product-specific information from 2D vaccine vial barcodes. We sought to examine the feasibility, performance, and potential limitations of scanning 2D barcodes on vaccine vials using 4 different mobile phones.

Methods: A unique barcode scanning app was developed for Android and iOS operating systems. The impact of 4 variables on the scan success rate, data accuracy, and time to scan were examined: barcode size, curvature, fading, and ambient lighting conditions. Two experimenters performed 4 trials 10 times each, amounting to a total of 2160 barcode scan attempts.

Results: Of the 1832 successful scans performed in this evaluation, zero produced incorrect data. Five-millimeter barcodes were the slowest to scan, although only by 0.5 seconds on average. Barcodes with up to 50% fading had a 100% success rate, but success rate deteriorated beyond 60% fading. Curved barcodes took longer to scan compared with flat, but success rate deterioration was only observed at a vial diameter of 10 mm. Light conditions did not affect success rate or scan time between 500 lux and 20 lux. Conditions below 20 lux impeded the device's ability to scan successfully. Variability in scan time was observed across devices in all trials performed.

Conclusions: 2D vaccine barcoding is possible using mobile devices and is successful under the majority of conditions examined. Manufacturers utilizing 2D barcodes should take into consideration the impact of factors that limit scan success rates. Future studies should evaluate the effect of mobile barcoding on workflow and vaccine administrator acceptance.

Introduction

As digital health infrastructure evolves, the inclusion of product-specific identifiers in electronic health records will become of greater importance. This is particularly relevant for immunization practice where lot numbers in patient records are essential for the evaluation and surveillance of vaccine safety and effectiveness at the product level. However, product-specific identifiers are often recorded by hand, resulting in missing or inaccurate information. Missing data are known to produce gaps

in communication between health care providers, increasing the potential for poor care coordination and medical errors [1]. Examination of children's immunization records reveals transcription errors (in some cases more than 10%), administration of look-alike or sound-alike products, sibling confusion, and repeat immunization [2].

Vaccine products that protect against the same diseases are not necessarily the same formulation. The differentiation between vaccine products in vaccination records is essential for evaluations of the safety and effectiveness of vaccines. In order to identify different vaccine products, two-dimensional (2D) barcodes are often printed on vaccine vials. The most commonly employed 2D barcode standard for vaccine vials is the DataMatrix. At 2-3 mm^2, DataMatrix barcodes can store up to 50 alphanumeric characters, making them capable of containing a Global Trade Item Number (GTIN), an expiration date, and a lot number, in an image small enough to be printed directly on unit-of-use product labels [3]. GTINs are identification numbers that are used to identify products all over the world. 2D barcode scanning has the potential to play an important role in automating the identification of vaccines such that they can be included in electronic health records efficiently and with fewer errors [4].

Barcode scanning of vaccine products is not widely implemented, although preliminary implementation pilots are positive, showing improvements in data completeness and reduction in data errors [5]. A time-motion study demonstrated that scanning 2D barcoded vaccines could reduce immunization documentation time by 36-39 seconds per dose [6]. Training requirements and process flow issues, access to and adoption of technology, and resistance to change are known barriers to the implementation of barcode scanning within health facilities [7].

Barcode scanning facilitated by mobile devices such as mobile phones could potentially increase the amount of vaccine product-specific information residing in digital health records by making barcode scanning more readily accessible to both health care providers and patients. Health care providers could use the mobile device they already own as a scanner instead of purchasing a handheld scanner. Additionally, health care providers working in remote areas where carrying a handheld scanner is not feasible would likely still be able to use their

mobile device to capture data. Enabling patients to capture their own product-specific records could also be beneficial, especially within immunization where parents are often responsible for maintaining their children's immunization data. Although it is unlikely that a parent would be given the vaccine vial to scan, a barcode could be provided to a patient on a vaccine information sheet, which the patient could scan to capture the information into a personal vaccination record app. The feasibility of mobile barcode scanning of vaccine vials and its limitations remain uncertain. Our objective in this study was to examine the feasibility, potential limitations, and variability in performance of scanning vaccine vial barcodes using mobile phones.

Methods

Objectives

We sought to determine whether mobile phones are capable of accurately scanning 2D vaccine barcodes. We specifically examined the impact of barcode size, curvature, fading, and lighting on the ability to successfully scan 2D barcodes, as well as how barcode scanning ability varies among different mobile devices.

Study Setting and Variables Examined

A mobile phone app was developed for iOS and Android platforms that scans barcodes and records whether the scan was successful within an allotted amount of time. The app was developed by programmers at the Ottawa Hospital Research Institute specifically to perform this study. The time to scan the barcode was also recorded. The app was loaded onto 4 different mobile phone devices that were state of the art in mid to late 2013: the iPhone 5, the Samsung Galaxy S4, the Nexus 5, and the Nexus 7 (Table 1) [8-11]. The mobile app developed was used to perform a validation study on the ability of the devices to scan the 2D barcodes under a variety of laboratory conditions.

The experiment was divided into 4 trials, each evaluating the effect of 1 variable (barcode size, fading, curvature, and ambient lighting) on the scannability of perfectly printed 2D DataMatrix barcodes as recorded by the app (Table 1). A series of barcode samples was produced for each trial and printed on standard printer paper using an ink-jet printer (Multimedia Appendix 1).

Table 1. Trial conditions.

Trial	Measuring	Size	Curvature	Fading	Ambient light	No. of scans
1	Size	Varied in 0.5-mm increments between 5 mm and 9 mm	Flat	None	500 lux	9
2	Fading	7 mm	Flat	Varied in 10% increments between 0% and 90%	500 lux	10
3	Curvature	7 mm	0 mm (flat), 10 mm, 15 mm, 17 mm	None	500 lux	4
4	Ambient light	7 mm	Flat	None	5, 20, 150, 500 lux	4

The study was performed in Ottawa, Canada. All trials were performed in a room with no natural light, with a light source

fixed at a specific illuminance. Illuminance is a measure of the quantity of light travelling past a surface and was measured

using a lux meter adjacent to the location where the vials were scanned. Preliminary experiments were performed before the validation study in order to characterize the range of values for each independent variable. A baseline value was determined for each variable at the value where the variable no longer had an effect on the ability of all devices to scan barcodes one hundred percent of the time. Holding these baseline values constant for 3 of the variables allowed us to isolate the effect of the fourth variable. The baseline values were determined by repeatedly attempting to scan barcodes while increasing the variable parameter until 10 successive scan attempts succeeded, for each of the 4 devices.

During all of the trials, except that which evaluated the ambient lighting variable, the illuminance was fixed as close to 500 lux as possible. An illuminance of 500 lux was established as the baseline illuminance in our preliminary experiments. The Canadian Occupational Health and Safety Regulations recommend illuminance levels of 1000 lux in examination and treatment rooms and 500 lux in other health care environments [12].

Size

To evaluate the impact of size on barcode scannability, the sample set consisted of a series of 9 barcodes decreasing in size by 0.5 mm from 9 mm to 5 mm. The value of 5 mm was chosen as the lower limit, as this is the barcode size present on single-dose vaccine syringes. A value of 7 mm was identified as the baseline size used to eliminate the effect of size on the other trials.

Fading

For the fading variable, the sample set consisted of a series of 10 barcodes. Fading was applied such that a barcode with 0% fading would be printed with full black color and a barcode with 100% fading would be invisible. The series used for the trial consisted of barcodes with increasing fading in increments of 10% from 0% to 90%. The upper limit was 90% as at 100% the barcode is not visible.

Curvature

To evaluate curvature, 4 barcodes with a uniform size of 7 mm were printed on adhesive paper and pressed onto vaccine vials with diameters of 10 mm, 15 mm, and 17 mm. These diameters correspond to those of the Sanofi Pasteur 0.5-mL-dose syringe, 1-mL-dose vial, and 0.5-mL-dose vial, respectively. A fourth barcode was printed on a flat surface. All other trials were performed entirely with flat barcodes.

Lighting

To evaluate ambient lighting, ideal 7-mm printed barcodes were scanned at 4 illuminance levels. A lux meter was used to measure the illuminance in the immediate area where the barcode was scanned. A dimmer was used to adjust the lighting in the room to specific light intensities, as may be experienced in different clinical settings. The 4 illuminance levels tested were 500, 150, 20, and 0 lux. The illuminance 20 lux was chosen because in our initial experiments scanning became difficult

around this point. Lastly, we chose 5 lux to simulate a near pitch-dark environment.

Study Procedure

The first page of the app allows the experimenters to select the trial they want to perform and to input their name (Multimedia Appendix 2, Screenshot 1). For all of the trials except the ambient lighting trial, the experimenter is required to enter the illuminance measurement at which the trial is being performed.

To complete a single trial, an experimenter scanned each barcode in the series once. Each trial was performed 10 times on each of the 4 devices, by 2 independent experimenters. The experimenters went through a training period where they each performed each trial twice to familiarize themselves with the scanning procedure. For all trials except the curvature trial, the paper containing the barcode was fastened to a flat surface. For the curvature trial, the vials to which the barcodes were adhered were fastened to the surface with the barcodes facing up.

When the experimenter begins a trial, a screen is shown that indicates how many scans are remaining in the trial and which barcode must be scanned next (Multimedia Appendix 2, Screenshot 2). When the user taps the scan button, the device must be in the user's hand, which must be resting on the table surface 8 inches to the right of the sample. Only after the scan button has been pressed may the user move the device to attempt to scan the barcode. This was done to eliminate disparities between scans due to the experimenter maintaining the exact position and height of the device, which promotes subsequent scans being far quicker than the initial scan.

Analysis

A scan is defined as successful when the scanner reads the correct code printed in the barcode within 10 seconds. We chose the limit of 10 seconds as we expect this to be the maximum amount of time a user would continue attempting to scan a vial without success [13]. In addition to the 4 variables mentioned above, we also looked at the percentage of scans that succeeded but returned an incorrect value. To measure this, the app checks whether each scan returns the value encoded in the barcodes in the sample and records this as part of the scan data (Table 2 in Multimedia Appendix 3). The app outputs the results of each trial as a comma-separated vector file. The interrater reliability was calculated between the 2 raters using the two-way random, single measure, intraclass correlation (ICC). The values for each subject compared in the ICC calculation are the scan success rates (out of 10) of each rater for each subject. A subject, for our purpose, is defined as the set of 10 scans that have the same variables: trial type, sequence number, and device.

Results

Accuracy of Barcode Scanning

Each experimenter performed all 4 trials 10 times each, amounting to a total number of 2160 barcode scan attempts. Out of the 1832 successful scans there were zero scans that registered as successful but produced incorrect data (Table 2).

Table 2. Total successful scans for all trial conditions and by device.

Trial condition	Total success rate, %	Successful scans by device, n			
		Nexus 7	iPhone 5	Nexus 5	Samsung Galaxy S4
Size, mm					
5	100.00	20	20	20	20
5.5	100.00	20	20	20	20
6	100.00	20	20	20	20
6.5	100.00	20	20	20	20
7	100.00	20	20	20	20
7.5	100.00	20	20	20	20
8	100.00	20	20	20	20
8.5	100.00	20	20	20	20
9	100.00	20	20	20	20
Fading, %					
0	100.00	20	20	20	20
10	100.00	20	20	20	20
20	100.00	20	20	20	20
30	100.00	20	20	20	20
40	100.00	20	20	20	20
50	100.00	20	20	20	20
60	77.50	14	16	13	19
70	1.25	0	0	0	1
80	0.00	0	0	0	0
90	0.00	0	0	0	0
Curvature, mm					
0	100.00	20	20	20	20
10	88.75	20	18	17	16
15	100.00	20	20	20	20
17	98.75	20	20	20	19
Illuminance, lux					
500	100.00	20	20	20	20
150	100.00	20	20	20	20
20	95.00	20	20	18	18
5	28.75	0	0	7	16

Interrater Reliability

The ICC between the 2 raters was observed to be .947 with a confidence interval of .921 to .964. Table 3 presents the number of successful scans per rater by device for each subject. The most significant deviation between the 2 raters occurs during the fading trial at 60% fading, where rater 1 is more successful on every device than rater 2 by approximately 180% on average. The second largest deviations occur in the curvature and lighting trials on the Samsung Galaxy S4 device, where in both cases rater 1 succeeds on every attempt and rater 2 succeeds on 86.25% of the attempts. For each trial, the point at which interrater reliability begins to deteriorate likely borders on the limitations of the technology for practical use.

XSL·FO

RenderX

Table 3. Total number of successful scans per device and rater.

Trial Condition	Successful Scans by Device and Rater							
	Nexus 7		iPhone 5		Nexus 5		Samsung Galaxy S4	
	Rater 1	Rater 2	Rater 1	Rater 2	Rater 1	Rater 2	Rater 1	Rater 2
Size, mm								
5	10	10	10	10	10	10	10	10
5.5	10	10	10	10	10	10	10	10
6	10	10	10	10	10	10	10	10
6.5	10	10	10	10	10	10	10	10
7	10	10	10	10	10	10	10	10
7.5	10	10	10	10	10	10	10	10
8	10	10	10	10	10	10	10	10
8.5	10	10	10	10	10	10	10	10
9	10	10	10	10	10	10	10	10
Fading, %								
0	10	10	10	10	10	10	10	10
0.1	10	10	10	10	10	10	10	10
0.2	10	10	10	10	10	10	10	10
0.3	10	10	10	10	10	10	10	10
0.4	10	10	10	10	10	10	10	10
0.5	10	10	10	10	10	10	10	10
0.6	10	4	10	6	9	4	10	9
0.7	0	0	0	0	0	0	1	0
0.8	0	0	0	0	0	0	0	0
0.9	0	0	0	0	0	0	0	0
Curvature, mm								
0	10	10	10	10	10	10	10	10
10	10	10	10	8	8	9	10	6
15	10	10	10	10	10	10	10	10
17	10	10	10	10	10	10	10	9
Illuminance, lux								
5	0	0	0	0	3	4	10	6
20	10	10	10	10	8	10	10	8
150	10	10	10	10	10	10	10	10
500	10	10	10	10	10	10	10	10

Impact of Size

The size trial yielded a 100% scan success rate, meaning that every scan attempt succeeded before timing out (Table 2). The results of the size trial indicated that 5-mm barcodes took longer to scan than any of the other barcode sizes we tested, although only by 0.5 seconds on average. Figure 1 shows the average scan time across all devices and experimenters for each barcode size tested.

When we examined how scan time differed among devices (Figure 1), the Samsung Galaxy S4 tended to have consistently longer scan times (0.55 seconds more) than the other devices. The iPhone consistently had the lowest average scan time.

XSL•FO
RenderX

Figure 1. Scan Time by Barcode Size and Device.
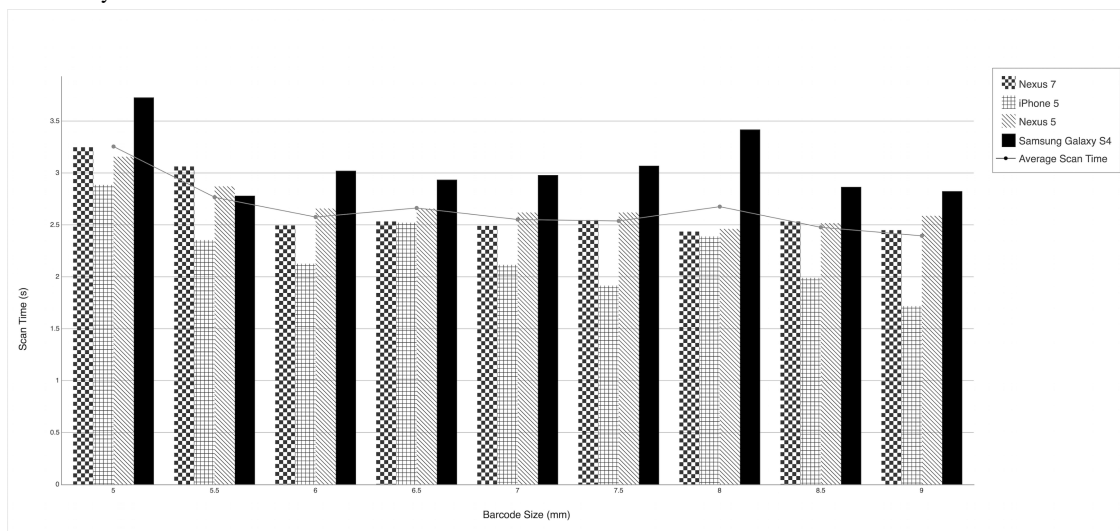
Impact of Fading

The results of the fading trial demonstrated that fading begins to affect scan time and overall scannability at 60% and becomes nearly impossible at 70%. Table 2 shows that up until 50% there is a 100% scan success rate. At 60% fading, however, the scan success rate decreases to ~78%.

The Samsung Galaxy S4 device exhibited superior scan time and scan success percentage and was the only device to scan the 70% faded barcode (Figure 2, Table 2). Up until the 60% faded barcode, the performance of all the other devices was mostly uniform.

Figure 2. Scan Time by Barcode Fading and Device.
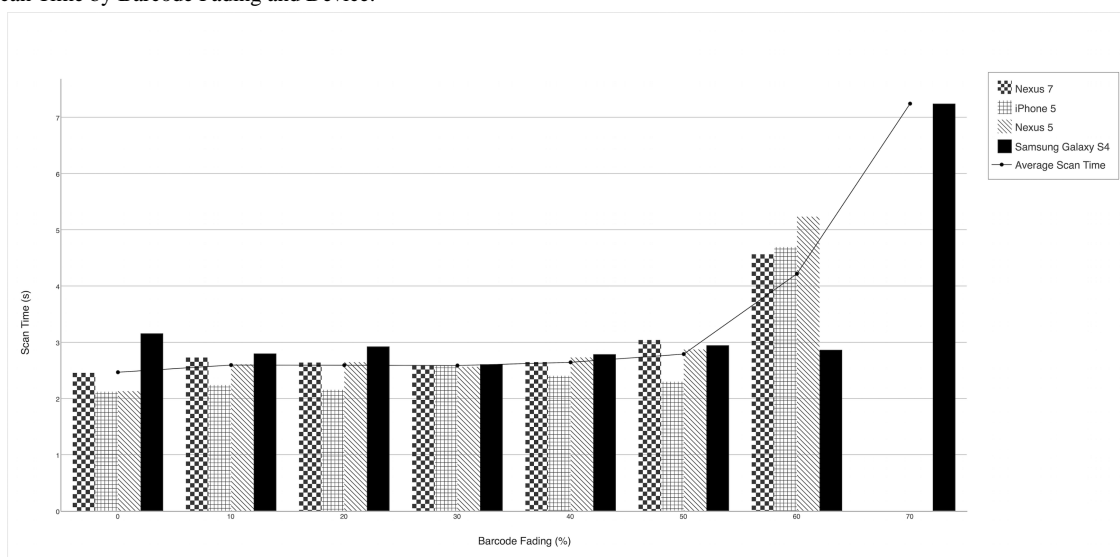
Impact of Curvature

The results of the curvature trial can be seen in Figure 3. The curvature with the smallest diameter (10 mm) saw the longest scan time and lowest scan success percentage at 3.5 seconds and 88%, respectively (Table 2). On average, all of the curved barcodes took longer to scan than flat barcodes; however, the

scan success rate remained at approximately 100% with the exception of the barcode with the 10-mm diameter of curvature. When examining scan time by device presented in Figure 3, the iPhone handled curvature better than any of the other devices with an average scan time approximately 1 full second less than the average. The Samsung Galaxy S4 had the longest scan time, with an average scan time 0.6 seconds higher than the average.

XSL•FO

RenderX

Figure 3. Scan Time by Barcode Curvature and Device.
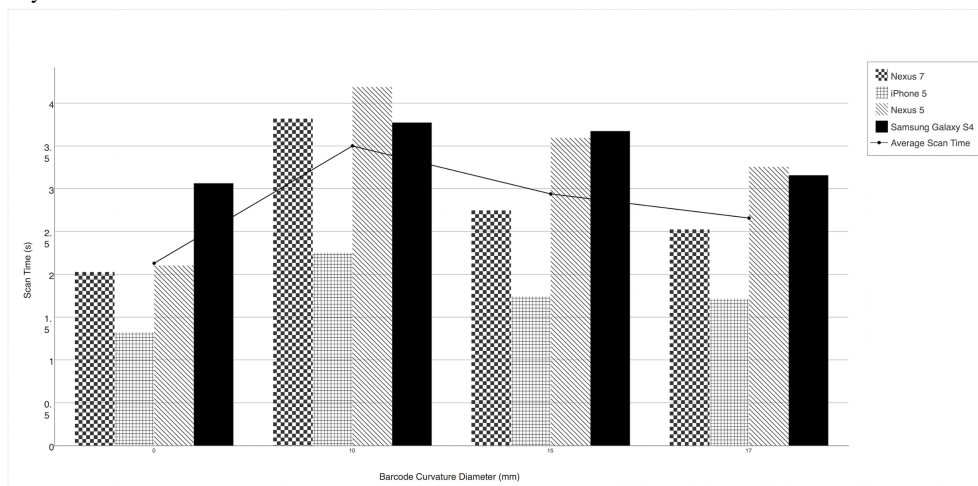
Impact of Lighting

Figure 4 shows the effect of varying the illuminance at the barcode on scan time. Scannability only degraded significantly once illuminance dropped below 20 lux (Table 2). The Samsung Galaxy S4 and Nexus 5 devices performed the best at the lowest lighting condition, whereas the iPhone and Nexus 7 were completely unable to scan the barcode at 5 lux. For the top 3 illuminance levels, the performance of the devices was mostly uniform with the iPhone exhibiting slightly shorter scan times, as it tended to do throughout the experiment (Figure 4).

Figure 4. Scan time by barcode illuminance and device.
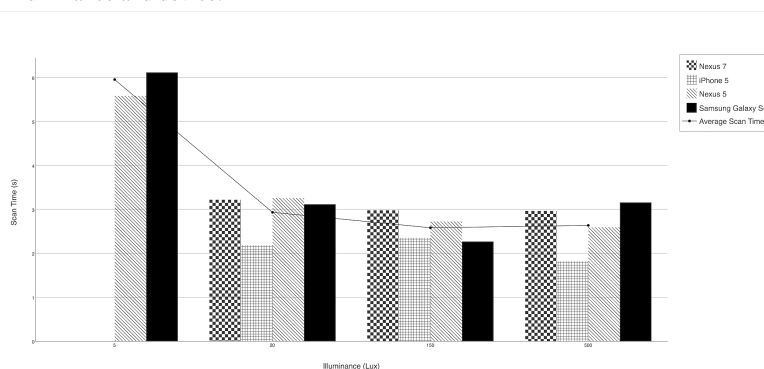
Discussion

We successfully developed a mechanism for scanning 2D barcodes using mobile devices. We had a 100% data accuracy rate for all barcodes successfully scanned. When examining factors potentially limiting 2D barcode scanning, our study found that, given ideally printed barcodes and using modern mobile phones, the following are true:

1. Barcodes as small as 5 mm can be scanned reliably. The average scan time could be marginally increased by using barcodes larger than 5 mm.

2. Scannability begins to decrease significantly when the barcode has faded past 50%.

3. Curvature begins to affect scannability between 10- and 15-mm diameters.

4. Illuminance begins to deplete scannability around 20 lux.

Our results suggest that modern mobile phones should be able to scan barcodes printed on vaccine vials and other packaging, assuming those barcodes are printed without errors, larger than or equal to 5 mm, and do not exhibit fading greater than 50%.

Performance was mostly uniform across all devices tested. It became evident that scan time was mostly dependent on the properties of the software program as opposed to the hardware properties of the device. For instance, the iPhone outperformed any other device on almost every trial when looking at scan time; however, this is most likely due to the settings of the auto-focus timer that is responsible for periodically adjusting the focus of the camera in order to refocus on the subject. The auto-focus timer on the Android devices was approximately 2 seconds, whereas the auto-focus timer on the iPhone was approximately 1.5 seconds. It is this disparity that is largely responsible for the faster scan times of the iPhone. Despite device variation, our average scan time results were similar to what was found by Pereira et al [14] who utilized scanners retailing for almost US $800 (PowerScan D8530 Handheld Scanner, Datalogic Mobile Inc).

Although it was to be expected, the absolute absence of data errors (successful scans yielding incorrect data) is an important outcome as it confirms the reliability of mobile barcode scanning with respect to data integrity. Data errors are possible, especially in a live setting where barcodes are not necessarily printed with perfect precision. In most jurisdictions, parents are still required

to maintain paper records of their children's immunizations and some surveys indicate that more than a quarter of these records are incomplete, contain data errors, or are lost completely [15]. Accurate tracking of an individual's immunization history is important in order to prevent duplicate immunizations and when proof of immunization is required for school, day care, and so on. Recording errors are common due to vaccine or patient-related human factors [16]. As individuals increasingly receive vaccines from multiple providers, it is important that individuals maintain accurate records as they will often be viewed as the single source of truth when it comes to their immunization status [17]. Research demonstrates that increasing responsibility and control over one's own health records, including immunization history, fosters greater engagement within the health care system and increases knowledge about personal health [18]. As registries in many jurisdictions underestimate coverage [19,20], any mechanism that empowers the individual to record and report immunization encounters has the potential to improve immunization programs.

Although this study focusses on vaccine vials, 2D DataMatrix barcodes are being used increasingly to identify other medications and medical devices. One application that could be of particular interest to the public would be the use of barcode scanning to include product expiry information in consumer mobile apps. A person who depends on an inhaler could use his or her mobile device to scan the 2D barcode and capture the expiry date into an app that will remind the user to renew the prescription before it expires.

Mobile barcode scanning, like other mobile technologies [21], could provide value to practitioners. This may include data entry efficiencies for product-specific information at the point of care, resulting in fewer information gaps [22]. When barcode scanning is not available, physicians and nurses must read the information off of the vaccine packaging or vial and enter the information by either manually typing or selecting choices from drop-down menus. Both of these methods have been shown to produce more errors in hospital electronic medical records compared with barcoding methods [23].

To the best of our knowledge there have been no other studies that evaluate the limits of mobile barcode scanning. This study benefits from its use of 2 experimenters. We observed high interrater reliability between the 2 experimenters, which suggests that our results are reproducible.

A limitation of this study is that we did not evaluate the ability of mobile devices to scan the barcodes on real vials. There are three potential problems: first, from observation we know that the barcodes on some vaccine vials are not printed with sufficient quality to permit scanning with mobile devices. Some barcodes are printed with defects and many have been shown to exhibit fading of the print [24]. It is likely that the print quality will have to improve before scanning with mobile devices is entirely feasible. Second, our study evaluates the 4 variables independently, whereas when scanning an actual vial there will be an interaction of variables. For instance, it is possible that in low-light environments, the threshold, at which a faded barcode becomes impossible to scan, is higher. Consequently, the study benefits from observing each variable

independently in that we can say definitively what impact each variable has on scannability and our results are not skewed by the imperfect print quality of the barcodes on actual vials. In addition, to thoroughly evaluate the interdependence of each variable would require hundreds of thousands of scans, which may not be feasible. Last, the surface on which the barcode is printed and the reflectiveness property of the material would likely have an impact on scannability as well and would need to be evaluated as a fifth variable.

Another limitation of this study was the scope of devices we tested. The devices we used were state of the art in mid-2013. There exists a wide range of devices of lower quality both in computing power and in camera quality. Including some of these lower-quality devices in the study may have given a better indication of the lower limit of scannability. Since 2013, mobile devices have improved considerably, allowing for a shift away from relying on desktop processors [25]. We have made the reasonable assumption that newer phones that have higher specifications than the devices we tested will have equivalent or improved scanning performance.

Immunization information systems and registries are critical to the success of immunization programs [26] and become more powerful when they include product-specific identifiers. Mobile barcode scanning could serve as a mechanism to increase the amount of product-specific information captured in these systems by lowering the barriers to entry of barcode scanning at point of care [17]. We propose that there are two further components to the evaluation of mobile barcode scanning as an approach to the automated identification of vaccine products. First, an evaluation of mobile barcode scanning performance using real vaccine vials will be important to determine whether barcodes on vaccine products are being printed with sufficient precision to facilitate barcode scanning. Second, a usability study is necessary to determine whether mobile barcode scanning by individuals with little training is feasible and advantageous to manual text entry or drop-down selections. Although these two evaluations could tell us that mobile devices could in fact be used to facilitate barcode scanning, the greatest obstacle to the use of this technology will be incentivizing individuals and practitioners to use it. A third evaluation should be performed to determine whether mobile barcode scanning can be integrated into the physician's or nurse's workflow when vaccinating a patient.

Conclusions

This study has demonstrated that accurate 2D vaccine barcode scanning by mobile devices is possible and can be successful under the majority of laboratory conditions we examined. Within the context of vaccine barcoding in Canada, our results suggest that modern mobile phones should be able to scan barcodes printed on vaccine vials and packaging, assuming those barcodes are printed without errors, larger than or equal to 5 mm, and do not exhibit fading greater than 50%. Barcode scanning has been demonstrated to have a positive effect on the quality of health records [23], and mobile barcode scanning makes the technology available to a far greater number of health care providers and individuals than would otherwise have access to handheld scanners. The need for product- and lot-specific information to

XSL·FO

RenderX

be captured when building immunization information systems suggests a need for increased access to barcode scanning, which this study suggests could be fulfilled using mobile devices. Mobile barcode scanning should be considered as an adjunct to barcode scanning using dedicated handheld scanners, and manufacturers utilizing barcodes should take into consideration the factors that limit scanning.

Acknowledgments

Authors' Contributions

All authors were involved in the development of ImmunizeCA, a pan-Canadian immunization mobile app.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Barcode samples.

[PDF File (Adobe PDF File), 196KB - jmir_v18i6e143_app1.pdf]

Multimedia Appendix 2

App Screenshots.

[PDF File (Adobe PDF File), 94KB - jmir_v18i6e143_app2.pdf]

Multimedia Appendix 3

Device specifications and data captured per scan.

[PDF File (Adobe PDF File), 30KB - jmir_v18i6e143_app3.pdf]

References

1. Ancker JS, Witteman HO, Hafeez B, Provencher T, Van de Graaf Mary, Wei E. The invisible work of personal health information management among people with multiple chronic conditions: qualitative interview study among patients and providers. J Med Internet Res 2015;17(6):e137 [FREE Full text] [doi: 10.2196/jmir.4381] [Medline: 26043709]
2. Wilton R, Pennisi AJ. Evaluating the accuracy of transcribed computer-stored immunization data. Pediatrics 1994 Dec;94(6 Pt 1):902-906. [Medline: 7971009]
3. O'Connor A, Haque S, Layton C. RTI International, Research Triangle Park, NC. 2012. Impact of a Two-Dimensional Barcode for Vaccine Production, Clinical Documentation, and Public Health Reporting and Tracking URL: http://www.cdc.gov/vaccines/programs/iis/activities/downloads/2d-barcode-trkg-rpt.pdf [accessed 2015-12-15] [WebCite Cache ID 6ez7jcwWP]
4. Simpson N. Healthcare Information and Management Systems Society. Advocacy White Paper: Bar coding for patient safety Abstract URL: http://www.himss.org/ResourceLibrary/ResourceDetail.aspx?ItemNumber=9351 [accessed 2016-02-02] [WebCite Cache ID 6ez94hlyZ]
5. Fierro L, Gaddis M, Kinney M. 2014. Implementation Pilot for Two-Dimensional (2D) Vaccine Barcode Utilization: Summary Report URL: http://www.cdc.gov/vaccines/programs/iis/2d-vaccine-barcodes/downloads/pilot-summary.pdf [accessed 2016-02-02] [WebCite Cache ID 6ez9ZRMH0]
6. O'Connor AC, Kennedy ED, Loomis RJ, Haque SN, Layton CM, Williams WW, et al. Prospective cost-benefit analysis of a two-dimensional barcode for vaccine production, clinical documentation, and public health reporting and tracking. Vaccine 2013 Jun 28;31(31):3179-3186. [doi: 10.1016/j.vaccine.2013.04.073] [Medline: 23664988]
7. Nanji KC, Cina J, Patel N, Churchill W, Gandhi TK, Poon EG. Overcoming barriers to the implementation of a pharmacy bar code scanning system for medication dispensing: a case study. J Am Med Inform Assoc 2009;16(5):645-650 [FREE Full text] [doi: 10.1197/jamia.M3107] [Medline: 19567797]
8. Thinnest, Lightest iPhone Ever Features All-New Aluminum Design, Stunning 4-Inch Retina Display, A6 Chip & Ultrafast Wireless. Apple Introduces iPhone 5 URL: http://www.apple.com/ca/pr/library/2012/09/12Apple-Introduces-iPhone-5.html [accessed 2016-03-24] [WebCite Cache ID 6foe0pIf0]

9. Galaxy S4 gets you closer to what matters in life, bringing your world together. New York, USA; 2013. Samsung Introduces the Galaxy S4- A Life Companion for a richer, simpler and fuller life URL: http://www.samsungmobilepress.com/2013/03/14/Samsung-Introduces-the-GALAXY-S-4----A-Life-Companion-for-a-richer,-simpler-and-fuller-life [accessed 2016-03-06] [WebCite Cache ID 6fodosJPG]

10. Slimmest and Fastest Nexus Phone Ever Now Available Online in Select Markets. Seoul, Korea; 2013. LG and Google Introduce Google Nexus 5 URL: http://www.lgnewsroom.com/2013/11/lg-and-google-introduce-google-nexus-5/ [accessed 2016-04-04] [WebCite Cache ID 6gWIdXchJ]

11. ASUS. 2013. ASUS and Google Announce New Nexus 7 Tablet URL: http://www.asus.com/us/News/9WqUAWjWE31jr4JD [accessed 2016-04-04] [WebCite Cache ID 6gWIlphOu]

12. Government of Canada. Canada Occupational Health and Safety Regulations URL: http://laws.justice.gc.ca/PDF/SOR-86-304.pdf [accessed 2016-02-02] [WebCite Cache ID 6ezAMgVbM]

13. Nielsen J. Nielsen Norman Group. 1993. Response Times: The 3 Important Limits URL: https://www.nngroup.com/articles/response-times-3-important-limits/ [accessed 2016-04-04] [WebCite Cache ID 6gWJQ3dtC]

14. Pereira JA, Quach S, Hamid JS, Heidebrecht CL, Quan SD, Nassif J, Public Health Agency of Canada/Canadian Institutes of Health Research Influenza Research Network (PCIRN) Vaccine Coverage Theme Group. Exploring the feasibility of integrating barcode scanning technology into vaccine inventory recording in seasonal influenza vaccination clinics. Vaccine 2012 Jan 17;30(4):794-802. [doi: 10.1016/j.vaccine.2011.11.043] [Medline: 22119585]

15. Public Health Agency of Canada (PHAC). Canada Communicable Disease Report (CCDR). Measuring Up: Results from the National Immunization Coverage Survey, 2002 URL: http://www.phac-aspc.gc.ca/publicat/ccdr-rmtc/04vol30/dr3005a-eng.php [accessed 2016-01-19] [WebCite Cache ID 6ezAbOVkg3035]

16. Bundy DG, Shore AD, Morlock LL, Miller MR. Pediatric vaccination errors: application of the "5 rights" framework to a national error reporting database. Vaccine 2009 Jun 12;27(29):3890-3896. [doi: 10.1016/j.vaccine.2009.04.005] [Medline: 19442422]

17. Wilson K, Atkinson K, Deeks S. Improving vaccine registries through mobile technologies: a vision for mobile enhanced Immunization information systems. Journal of the American Medical Informatics Association 2016;23(1):207-211. [Medline: 26078414]

18. Zarcadoolas C, Vaughon WL, Czaja SJ, Levy J, Rockoff ML. Consumers' perceptions of patient-accessible electronic medical records. J Med Internet Res 2013;15(8):e168 [FREE Full text] [doi: 10.2196/jmir.2507] [Medline: 23978618]

19. Braeckman T, Lernout T, Top G, Paeps A, Roelants M, Hoppenbrouwers K, et al. Assessing vaccination coverage in infants, survey studies versus the Flemish immunisation register: achieving the best of both worlds. Vaccine 2014 Jan 9;32(3):345-349. [doi: 10.1016/j.vaccine.2013.11.041] [Medline: 24269616]

20. Breva LP, Domingo JD, Martínez Beneito Miguel Ángel, Barberà JP. Vaccine coverage estimation using a computerized vaccination registry with potential underreporting and a seroprevalence study. Vaccine 2015 Apr 27;33(18):2183-2188. [doi: 10.1016/j.vaccine.2015.02.048] [Medline: 25748335]

21. Prgomet M, Georgiou A, Westbrook JI. The impact of mobile handheld technology on hospital physicians' work practices and patient care: a systematic review. J Am Med Inform Assoc 2009;16(6):792-801 [FREE Full text] [doi: 10.1197/jamia.M3215] [Medline: 19717793]

22. Korosec L, Balenko K, Hagens S. Impact of information technology on information gaps in canadian ambulatory care encounters. JMIR Med Inform 2015;3(1):e1 [FREE Full text] [doi: 10.2196/medinform.4066] [Medline: 25595130]

23. Pereira JA, Quach S, Hamid JS, Quan SD, Diniz AJ, Van ER, Public Health Agency of Canada/Canadian Institutes of Health Research Influenza Research Network (PCIRN) Program DeliveryEvaluation Group. The integration of barcode scanning technology into Canadian public health immunization settings. Vaccine 2014 May 13;32(23):2748-2755 [FREE Full text] [doi: 10.1016/j.vaccine.2013.11.015] [Medline: 24252700]

24. Glaxosmithkline. 2014. Potential fading of lot-specific information on syringe and vials of all GSK vaccine products URL: http://www.fda.gov/downloads/BiologicsBloodVaccines/SafetyAvailability/UCM418974.pdf [accessed 2015-08-07] [WebCite Cache ID 6ezAxT7JL2]

25. Bonnington C. Wired Magazine. In less than two years, a smartphone could be your only computer URL: http://www.wired.com/2015/02/smartphone-only-computer/ [accessed 2016-02-02] [WebCite Cache ID 6ezBBMMfa]

26. Johansen K, Lopalco PL, Giesecke J. Immunisation registers--important for vaccinated individuals, vaccinators and public health. Euro Surveill 2012;17(16):e1 [FREE Full text] [Medline: 22551460]

Abbreviations

2D: two-dimensional
GTIN: Global Trade Item Number
ICC: intraclass correlation

XSL•FO
RenderX

<u>Original Paper</u>

Technology-Based Innovations to Foster Personalized Healthy Lifestyles and Well-Being: A Targeted Review

Emmanouil G Spanakis[1], BSc, MSc, PhD; Silvina Santana[2,3], PhD; Manolis Tsiknakis[1,4], BSc, MSc, PhD; Kostas Marias[1]; Vangelis Sakkalis[1]; António Teixeira[2,5]; Joris H Janssen[6,7]; Henri de Jong[8]; Chariklia Tziraki[9], BSc, MD, PhD

[1]Computational BioMedicine Laboratory (CBML), Institute of Computer Science (ICS), Foundation for Research and Technology (FORTH), Heraklion, Greece

[2]Institute of Electronics Engineering and Telematics of Aveiro (IEETA), University of Aveiro, Aveiro, Portugal

[3]Department of Economics, Management and Industrial Engineering, University of Aveiro, Aveiro, Portugal

[4]Assoc. Professor, Department of Informatics Engineering, Technological Educational Institute of Crete, Heraklion, Greece

[5]Department of Electronics, Telecommunications & Informatics, University of Aveiro, Aveiro, Portugal

[6]Sense Health, Rotterdam, Netherlands

[7]Department of Communication, Stanford University, San Francisco, CA, United States

[8]ASK Community Systems GmbH, Bad Schwalbach, Germany

[9]Association of Community Elders' Clubs (MELABEV), Jerusalem, Israel

Corresponding Author:
Emmanouil G Spanakis, BSc, MSc, PhD
Computational BioMedicine Laboratory (CBML)
Institute of Computer Science (ICS)
Foundation for Research and Technology (FORTH)
N. Plastira 100 Vassilika Vouton, Heraklion, Crete, Greece
Heraklion, GR-700 13
Greece
Phone: 30 2810391446
Fax: 30 2810 391428
Email: spanakis@ics.forth.gr

Abstract

Background: New community-based arrangements and novel technologies can empower individuals to be active participants in their health maintenance, enabling people to control and self-regulate their health and wellness and make better health- and lifestyle-related decisions. Mobile sensing technology and health systems responsive to individual profiles combined with cloud computing can expand innovation for new types of interoperable services that are consumer-oriented and community-based. This could fuel a paradigm shift in the way health care can be, or should be, provided and received, while lessening the burden on exhausted health and social care systems.

Objective: Our goal is to identify and discuss the main scientific and engineering challenges that need to be successfully addressed in delivering state-of-the-art, ubiquitous eHealth and mHealth services, including citizen-centered wellness management services, and reposition their role and potential within a broader context of diverse sociotechnical drivers, agents, and stakeholders.

Methods: We review the state-of-the-art relevant to the development and implementation of eHealth and mHealth services in critical domains. We identify and discuss scientific, engineering, and implementation-related challenges that need to be overcome to move research, development, and the market forward.

Results: Several important advances have been identified in the fields of systems for personalized health monitoring, such as smartphone platforms and intelligent ubiquitous services. Sensors embedded in smartphones and clothes are making the unobtrusive recognition of physical activity, behavior, and lifestyle possible, and thus the deployment of platforms for health assistance and citizen empowerment. Similarly, significant advances are observed in the domain of infrastructure supporting services. Still, many technical problems remain to be solved, combined with no less challenging issues related to security, privacy, trust, and organizational dynamics.

Conclusions: Delivering innovative ubiquitous eHealth and mHealth services, including citizen-centered wellness and lifestyle management services, goes well beyond the development of technical solutions. For the large-scale information and communication technology-supported adoption of healthier lifestyles to take place, crucial innovations are needed in the process of making and deploying usable empowering end-user services that are trusted and user-acceptable. Such innovations require multidomain, multilevel, transdisciplinary work, grounded in theory but driven by citizens' and health care professionals' needs, expectations, and capabilities and matched by business ability to bring innovation to the market.

Introduction

Non-communicable chronic diseases (NCD), primarily cardiovascular diseases, cancers, chronic respiratory diseases, and diabetes, are responsible for 63% of all deaths worldwide (36 million out of 57 million global deaths). More than 9 million of all deaths attributed to chronic diseases actually occur before the age of 60 [1]. The health and economic burden ensuing from the rising prevalence of NCDs, along with shrinking health care budgets in the face of an aging population, are major factors stifling innovation and employment opportunities [2].

The body of research is clearly directing policy to a paradigm shift, away from traditional disease treatment towards a person-centered individualized coproduction of health [3], promotion of healthier behaviors, and better coordination and management of health care [4]. Most NCDs could be better controlled and even prevented through healthier lifestyle choices made across the lifespan [5,6]. There is an urgent need for a new paradigm in health care systems [7].

Technological innovations may be key to tackling the next decade's challenge of how to use our social and economic capital to empower and motivate individuals to engage in healthier personal lifestyle choices [8]. Information and communication technologies (ICT) can allow for a bottom-up approach. They can be used to build person-centered and community-based health services for empowering individuals with knowledge, empathic support, security, and trust [9,10] that would motivate them to choose and sustain healthier daily lifestyles [11] and well-being.

Well-being is a widely used term encompassing various constructs [12] and addressed by several theoretical models [13] as shown in Table 1 [14-23]. The ecological model (see Figure 1) toward healthier lifestyles, well-being, and wellness takes into account multiple predisposing factors, requirements, and barriers such as intrapersonal variables [24] (eg, personality, health beliefs, knowledge, attitudes, and skills), interpersonal processes and their likely interactions with genetics, as well as community and macro/public policy levels factors. The arrow in Figure 1 extending across the four levels suggests that factors or barriers extend into and interact across various levels.

Table 1. Dimensions and determinants of wellness according to models in the literature [13].

Physical	Emotional physiological	Social	Intellectual	Spiritual	Occupational	Environmental	Reference
X	X	X	X	X	X		[14]
X	X	X	X	X	X	X	[15]
X	X	X	X	X	X		[16]
X		X	X				[17]
X	X	X	X	X	X	X	[18]
X	X	X		X	X	X	[18]
X	X	X	X	X			[19]
X	X	X	X	X	X		[20]
X	X	X	X	X		X	[21]
X	X					X	[22]
X	X	X				X	[23]

Research has shown that interventions for change are likely to take place, not in the traditional health care system, but rather in what sociologists have labeled "enabling spaces" [25]. Virtual or spatial enabling spaces provide the opportunity for "mingling, observing, and lingering" where peer learning and teaching is actualized within a community. The goal of this work is to identify and discuss key innovations that may foster healthy lifestyle choices and well-being [26-31].

Figure 1. An ecological model of factors influencing life style behaviors [24].
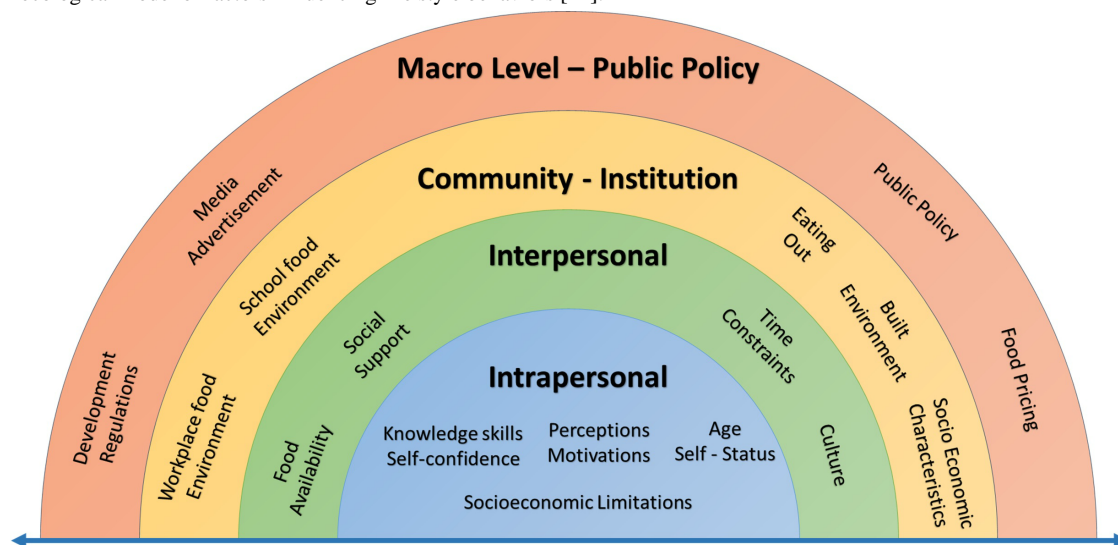
Methods

We follow a bottom-up approach for the review of state-of-the-art and beyond state-of-the-art theory and practice on personalized technology and innovations for empowering individuals to self-perceived well-being and healthy lifestyles. The question we are trying to answer is the following: What are the main sociotechnical challenges that need to be successfully addressed in delivering state-of-the-art, ubiquitous eHealth and mHealth services in the 21st century, including citizen-centered wellness management services for the large scale ICT-supported adoption of healthier lifestyles?

Our goal is to generate evidence on community-based, citizen-centered interventions and review key enabling technologies that can support research and bridge the gap between ICT health and social care systems [32]. We do so by critically reviewing selected state-of-the-art work in the field in order to identify some of the areas in need of further research. We present key concepts and themes that would contribute to consolidating the body of knowledge and propel further developments on the multidisciplinary sociotechnical aspects of health-related data collection, modeling, representation, and unification in a complex, person-centered ecosystem. We emphasize these innovations that are, in our opinion, necessary in order to sustain personalized healthy lifestyles and well-being for the public.

In more detail, this paper is structured as follows. First, we present the state-of-the-art technologies in the areas of personalized health monitoring systems, activity profiling, lifestyle capturing, and infrastructure supporting services, highlighting challenges in development and deployment of such technological solutions. Second, from the stakeholders' standpoint, we present a vision on specific key concepts that would channel a leap forward, that is, person-centered ICT-based innovations, ushering in a new generation of services enabling changes at the personal level and a paradigmatic shift in health and social care systems. These technologies could be used to achieve personalized healthy lifestyles, wellness, and

well-being through individual empowerment and engagement and effective self-management. We highlight and discuss specific innovations that would be able to support personalized healthy lifestyles and well-being for the public. Third, we discuss important drawbacks and pitfalls that must be addressed for the effective use of high-quality technological solutions that are able to serve people through secure, safe, and trustable innovative services to take place.

Results

State-of-the-Art Approaches, Techniques, and Technologies

Personalized health systems and pervasive mobile monitoring can enable sensing, mining, and learning of human behaviors and intentions. Examples include personalized mobile information delivery, context aware social networking, device and environment customization, serious games and entertainment, education, safety, and mobile business [33]. Their potential in the domain of lifestyle changes towards wellness, well-being, and self-management of chronic diseases, at the individual, organizational, and community levels is enormous, as can be seen from the following examples.

Personalized Health Monitoring Systems

mHealth and Smartphone Platforms

Smartphones are rapidly becoming the central computing, sensing (large number of sensors embedded), and communication platform for deploying personalized wellness mobile apps and services [34-36]. Smartphone capabilities include two types of sensing technologies: hardware-based sensors (physically present in the device) and software-based sensors (virtual sensors fusing multiple hardware sensors' data). These include accelerometers, global positioning system (GPS), digital barometer, microphone, camera, ambient light sensor, digital compass (Magnetometer), assisted GPS, proximity sensor, near field communication, global navigation satellite system, finger print reader, and many more [37].

Existing research [37-40] has evaluated smartphones as a method for delivering key components of established and empirically validated behavioral weight loss treatment, with an emphasis on adherence to self-monitoring. Results show that smartphones can be advantageous for optimizing adherence to self-monitoring and to the inclusion of behavioral strategies for evidence-based interventions. They have also been used to track social interaction, sleep, and physical activities, and to provide intelligent feedback promoting better health and well-being [41]. Currently, there are also several systems allowing developers to build apps and services on top of a smartphone's sensing capabilities [42-44].

Intelligent, Ubiquitous, and Smart Applications

Sensor-enabled mobile phones have the potential to collect *in situ* continuous sensor data that can dramatically change the way health and wellness are assessed and monitored, as well as how self-management of health conditions is made and care and treatment are delivered. New classes of applications are being explored both in academic- and industry-based research centers.

Advances in human-centric sensing are being fueled by the combination of sensor data and classification models to recognize human activities [45] and environmental context [46]. These apps correlate sensing information with personal health data and encourage users to be physically active and meet their related goals. Research is being conducted on using unobtrusive monitoring technology to study how mood changes are correlated with both social interactions and non-sedentary work style [47]. Examples exist on studying smoking habits and cessation, but the existing systems have low levels of adherence to key guidelines [48]. We found that only a few apps provide recommendations to the user for proven treatments such as pharmacotherapy, counseling, and/or a quit line.

A major challenge is to develop innovative personalized technologies that can help individuals maintain healthy lifestyles and wellness by keeping track of their everyday activities. Today, we have a plethora of apps and services supported by modern smartphones and mobile apps using embedded or external sensing devices. Most of these smart well-being tracking systems focus on capturing physical activity, fitness, and sleep patterns [43-48]. Major smartphone manufacturers include in their app suite dedicated proprietary software to allow people to self-monitor their health consistently, by logging and checking exercise, activity, sleep, food intake, and heart rate, among others [49-51]. All these apps are able to connect with specialized clock-like body-worn devices [52-54], allowing users to constantly record their activity and receive live feedback in real-time mode even if they do not carry their smartphone with them.

A recent study analyzed the content of many popular free apps related to physical activity and compared them against existing guidelines and fitness principles already established by the American College of Sports Medicine (ACSM) [55]. Results show that very few are evidence-based and respect the guidelines for aerobic activity, strength/resistance training, and flexibility, set forth by ACSM. The authors clearly identified a gap in safely and effectively using evidence-based apps to start a physical routine program, develop fitness, and lose weight. Nearly all the apps, although technically well designed, did not meet the basic recommendations of ACSM for exercise prescription, and therefore, would not be suitable for beginning exercisers. Thus, users are advised to select apps with extreme caution. Use of mobile technologies may have the potential to transform care delivery across populations and within individuals over time. However, such devices and/or services may need to be tailored to meet the specific patient and doctor needs [56,57].

Personalized Activity Profiling and Lifestyle Capturing

Human Activity and Lifestyle Recognition

Human activity recognition is a challenging problem for context-aware systems and apps. Research in this field has mainly adopted techniques based on supervised learning algorithms, but these systems suffer from scalability issues with respect to number of processed activities and richness of contextual data [58]. Efforts have been made toward the development of a unified framework for activity-recognition behavior-based analysis, and action prediction for the daily routine of people. A novel approach for enhanced classification of activity recognition data includes a game-like app used to reward physical activity and encourage healthy lifestyle choices [59,60]. Increasingly, context-aware systems are focusing on multiple domains besides physical activity, such as mental and social well-being [61].

Recent research has shown that emotions can now be recognized more accurately by machine learning algorithms than by laypersons [62]. Consequently, approaches for the automated recognition of different aspects of our social lives, like empathy, dominance, and non-verbal behavior have progressed rapidly [63-65]. Bringing these novel technologies into health apps will likely increase their effectiveness by further tracking important mental health parameters and personalizing services based on them.

Personalized Health Assistance and Interaction Support Platforms

In recent years, there have been many developments in the area of natural user interfaces, from touch screens to assistants integrating dialogue capabilities. Intelligent assistants that interact with users via conversational natural spoken language can provide them with meaningful and easily understandable information and advices (eg, about their prescribed medications) [65-68]. User-oriented, intuitive interaction is necessary in order to overcome the barriers of app acceptance. Special attention should be given to specific groups of people, like chronic patients and the elderly. Cardiac is such a prototype for a conversational assistant for chronic heart failure patients by using natural spoken dialogue [69]. Its helps patients managing their treatment and monitoring their health by using natural spoken dialogue over the telephone or with in-home systems to conduct regular checkups to collect relevant information on their condition. Similar research has focused on developing appropriate technical solutions, such as fission, information output arrangement, and organization modules for multimodal systems [70,71].

Infrastructure Supporting Services

Cloud for eHealth

Large amounts of data currently sit in different silos within health and social care systems. If these data were to be released in an appropriate manner and used effectively, they could transform the way care is provided [71]. Bringing together and correlating data among different and heterogeneous data sources would allow inference of new knowledge [72,73]. Cloud computing is emerging as a model for enabling convenient, on-demand network access to a shared pool of configurable resources that can be rapidly provisioned and released with minimal management effort or service provider interaction [74]. Institutions and medical professionals who frequently do not have enough storage and computing resources can manage their biomedical information through apps built on these type of services, accessing advanced computing infrastructures that they could not afford otherwise [75-78]. While many companies, like Google, IBM, Amazon, and Microsoft were early adopters of cloud computing, its application to biomedicine has only recently been proposed, mainly for bioinformatics applications.

Nevertheless, cloud computing is also benefiting the health care sector and the wellness management domain [77]. It can provide effective, on-demand, high elasticity access services to citizens from anywhere at any time. Data sharing on the cloud is quickly becoming vital for organizations and social users alike. In a survey by InformationWeek [79], nearly all organizations shared their data in one way or another with 74% sharing data with customers and 64% sharing data with suppliers.

The benefits that motivate organizations to move towards this direction include higher productivity and better time management. Health care providers are willing to store and share electronic medical records via the cloud and hence remove the geographical dependence between health care provider and patient [80]. Sarathy and Muralidhar [81] reviewed the impact of the Internet on data sharing across many different organizations such as government agencies and businesses. Butler [81] describes the issues of data sharing on the Internet where sharing information can allow users to infer details about users. Feldman and Patel et al [82] discuss the important benefit of data sharing in terms of public health, in particular for education and professional development.

However, the cloud is susceptible to privacy and security attacks, many of which occur from within the cloud providers themselves as they have direct access to stored data [83]. Malicious insiders represent one of the major issues affecting the cloud [84,85]. There is considerable work on protecting data from privacy and security attacks. The National Institute of Standards and Technology has developed guidelines to help consumers protect their data on the cloud [86]. Encrypting data before storage is an effective way to prevent unauthorized users from accessing sensitive data [87]. However, plain encryption techniques are not enough, especially when considering the scenario of sharing data among a large group of users [88].

Standards for Health and Interoperability

With the adoption of electronic health services, an opportunity exists for the creation of longitudinal health records that span

many decades and aggregate data from multiple health care organizations' source for delivering care. The need for standards for the representation and exchange of medical information becomes apparent. As defined by the Health Information Management Systems Society, an electronic health record (EHR) "is a longitudinal electronic record of patient health information generated by one or more encounters in any care delivery setting." According to the report of the National Institutes of Health (National Center for Research Resources) on EHRs, three main organizations create standards related to EHRs: the Health Level Seven (HL7), the Comité Européen de Normalisation - Technical Committee (CEN TC) 215, and the American Society for Testing and Materials (ASTM). HL7, which operates in the United States, develops the most widely used health care related electronic data exchange standards in North America. CEN TC 215, which operates in 19 European member states, is the preeminent organization developing health care information technology (IT) standards in Europe. Both HL7 and CEN collaborate with the ASTM, which is mainly used by commercial laboratory vendors. EHRs today use both technical and clinical standards. However, EHR vendors have implemented only some standards, while having a great deal of variation in their implementations, which results in systems that cannot interoperate and for which secondary use of data, for example, research and epidemiology, is difficult. Current EHR systems, due to their evolution over time, are often just an electronic representation of the previously used paper records. They are highly idiosyncratic, vendor-specific realizations of patient record subsets. They adopt few, if any, health information standards and very rarely accommodate controlled terminologies where they might be sensible.

A variety of health care communication standards has also been developed during the last decade. Their goal is to improve the interoperability and the connectivity among devices, computerized systems, tablets, smartphones, and health information systems, standardizing the content and structure of the information exchanged. Within the EU region, qualified medical devices and software depend on whether or not the device or software falls within the scope of the Medical Devices Directive (ie, 2007/47/EC). Currently, the directive states that software can indeed be qualified as a medical device, but unfortunately it does not specify what exact kind of software will meet the medical device definition per se.

The Continua Health Alliance guidelines [89] describe a set of standards to allow vendors, solution developers, and sharers of various types of health-related information to easily share their data. Some of the communication protocols supported by these guidelines are Bluetooth, ZigBee, USB, Wi-Fi, and Li-Fi. There are also tools available allowing vendors, programmers, and engineers to build on this approach. These resources are important in order to realize the communication interoperability recommended by the Continua Guidelines. Such examples are the Wipro Continua Toolkit (which enables medical devices to become compliant with the Continua specified protocol, ie IEEE 11073-XXXXX, and contains a Wipro Continua Agent and a Wipro Continua Manager), Stollman BlueHDP (health device profile)+USB dongle [90], Toshiba Bluetooth HDP stack and

application programming interface, and the Advanced and Adaptive Network Technology (ANT+) [91].

Service-Oriented Architecture and Service-Oriented Device Architectures

People do not experience technology. They experience services, made available by technology [92], or that should be the case. However, due to the heterogeneity in hardware and communication interfaces, interoperability is still a major concern in today's distributed architectures. This concern is particularly addressed by Service-Oriented Architecture (SOA) [93,94]. SOA implementations focus on principles such as separation of business logic from the underlying technology, efficient use and reuse of resources, compliance to standards, granularity and modularity, delivery, and monitoring of services. Service-Oriented Device Architecture (SODA) resulting from an adaptation of SOA intends to provide some level of abstraction to physical devices, simplifying external access [95-97].

Semantic Web/Linked Data and Ontological Approaches

The Semantic Web aims at a machine interpretable view of the World Wide Web. It focuses on structuring the Web based on content categorization to improve aspects such as automatic discovery, composition, invocation, and interoperation of services. The Semantic Web stack includes a number of technologies and standards aimed at this purpose (eg, Resource Description Framework or Web Ontology Language (OWL), SPARQL).

In the domain of health and wellness, a state-of-the-art example using semantic description is the GoPubMed [98] portal—a knowledge search engine for biomedical texts. At its core, GoPubMed uses two ontologies, GeneOntology [99] and MeSH [100] Ontology, but also allows users to create a custom ontology for specific searches. Another state-of-the-art example is NextBio, an ontology-based semantic framework based on gene, tissue, and disease ontological representation [101]. The use of ontology-based information extraction is explored in systems such as MedInX to allow semantic search of medical information originally buried in unstructured written text [102]. This approach is applicable both to private and public data, such as the one available on the Web.

One key problem with working with ontologies is that application ontologies tend to jeopardize the aim of ontology-based semantic interoperability to work with generalized references in the given domain. With respect to the biomedical arena, a suite of quality-checked, interoperable domain ontologies is being developed [103]. Yet, we need easy-to-use techniques to create application ontologies from the semantic information stored within those domain ontologies. This is a transformation of the bench-to-bed problem, since we have to bridge the gap between highly theoretical ontological representations to functionality-oriented ontologies.

Innovations to Support Healthy Life Styles and Well-Being

In this section, we discuss person-centered developments for the adoption of healthier lifestyles that are beyond the state-of-the-art. We emphasize the perspective of overarching ICT-supported innovations to influence the way health is maintained. Acquire, profile, represent, persuade, assess, manage, connect, and respond are some of the essential capabilities/functions/services needed to create innovative user-centered approaches.

A change toward healthier lifestyles is both attainable and urgent, given the current situation in most western world economies. Research shows that ICT resources and health behavior interventions can be brought together to make it feasible and effective, improve health outcomes, and bring down costs [104]. Behavior aware computing and pervasive technologies include novel patterns and activities relevant to individuals' health and well-being. Still, the problem of identifying behaviors from an individual's own annotated data, collected by multiple sources (eg, smartphones, sensors, and smart clothes) remains [105]. Care and most health-related practices continue to be caregiver-centered, despite recent efforts [106].

In order to go beyond the state-of-the-art, the user must be brought in as an active participant to care processes and related decision-making processes. To allow informed decision making, we need to advance intelligent user profiling and understandable instruction information to users. We need to enable interaction with Internet-of-Things (IoT) resources and services to retrieve and transmit information to users in simple everyday devices, such as mobile phones, smart watches, and home appliances [107]. This implies advancement in technologies and systems able to continuously track health-related data, despite coverage and communication means.

Healthy lifestyle change is attainable but not always easy to achieve, and sustaining behaviors, including use of ICT-based health-related programs, might be particularly challenging. Crucial evolution is needed in the process of making the essential information available and usable for the users, helping a person make a decision, and developing and deploying usable empowering end-user services that are accepted and enjoyed [108].

Acquiring User's Activity and Behavior In Digital Form

Many techniques have been proposed to automatically recognize human activities. Most important approaches use either statistical or symbolic reasoning. However, to date, these methods have mainly been considered separately. Proposed statistical activity recognition techniques differ on the kind and number of used sensors, activities considered, learning algorithms adopted, and many other parameters. One research direction focuses on using video with sound, image, and scene recognition software [109]. Other activity recognition techniques are based on data acquired from body-worn sensors (eg, motion tracking and inertial sensors, cardio frequency meters) and on the application of statistical learning methods. Early attempts in this field were mainly based on the use of data acquired from multiple body-worn accelerometers [110].

One of the main limitations of these early systems was the fact that they did not consider contextual information (eg, current location, environmental conditions, and surrounding objects)

XSL•FO

RenderX

that could be used to derive user activity. On the other hand, the recognition of complex activities like social ones (eg, work meetings, friendly chat) is particularly challenging and is hard to achieve by the use of solely statistical methods. Indeed, complex activities can be better recognized by considering constraints and relationships among context data that can be neither directly acquired from sensors nor derived through statistical reasoning alone.

As a result of these constraints, there is a growing interest in the use of ontology-based techniques to automatically recognize complex context data such as human activities. While most activity recognition systems rely on data-driven approaches, the use of knowledge-driven techniques is gaining increasing interest. Research in this field has mainly focused on the use of ontologies to specify the semantics of activities and ontological reasoning to recognize them based on context information. In particular, in the area of pervasive computing, OWL DL has been used to build activity ontologies and to recognize activities based on context data [111,112]. The ontological approach to activity modeling consists of a knowledge engineering task to define the formal semantics of human activities by means of the operators of the ontological language. Ontological reasoning is used to recognize that a user is performing a certain activity starting from some facts (eg, sensor data, location of persons and objects, properties of actors involved). Previous research has led to the definition of an OWL DL ontology for the activity recognition domain called COSAR [113], which is published on the PalSPOT [114] project website. This particular ontology definition was used to refine the predictions of statistical activity recognition systems by means of symbolic reasoning. The experimental evaluation of the effectiveness of this ontology-based activity recognition service, using a dataset collected in a smart-home setting, revealed the importance of including temporal reasoning in ontological techniques to effectively recognize activities.

Intelligent User Profiling for Healthy Lifestyles and Wellness

Intelligent user profiling [115] enables the collection of information from different sources to construct individual profiling models, with the objective of optimizing information delivery from health professionals to individuals in a personalized empowering environment. The most common user profile contents are the following: short/long-term user interests; knowledge, background, and skills; goals and user objectives or purpose with respect to the application; behavior; interaction preferences; and individual characteristics [116-119]. One such model is the OCEAN model [119-121].

User context is of great importance in the area of health, healthy lifestyles, and wellness management to characterize a situation of an entity [122]. Profiling information representation is following keyword-based models where interests are represented by weighted vectors or keywords [123]. Weights usually represent the relevance of the word for the user or within the topic. Another way to represent user interests is through topic hierarchies. Goals or intentions can be represented in different ways, either based on multicriteria analysis techniques [115] or Bayesian networks [124,125]. In this representation, nodes

represent user tasks and arcs represent probabilistic dependencies between tasks. Given evidence of a task performed by the user, the system can infer the next most probable task and hence the user's goal.

To obtain and build a user profile, the information can be provided explicitly by the user or be obtained through implicit observation of user actions. The simplest way of obtaining information about users is through the data they submit via forms or other user interfaces provided for this purpose. Especially for patients, their profile information is commonly assessed by patient-reported outcome measures (PROMs) including health-related quality of life information. PROMs can be defined as "reports coming directly from patients about how they function or feel in relation to a health condition and its therapy, without interpretation of the patients' responses by physician or anyone else" [126]. These instruments embrace a broad range of health dimensions such as physical, psychological, and social functioning [127]. On the other hand, in order to implicitly collect information about user's actions, their actions should be logged and patterns should be discovered using data mining, information retrieval, or machine learning techniques [128-131]. However, to be able to discover patterns, the user behavior should be repetitive, and the behavior observed should be different for different users. When no information about a user is available, a stereotype can be used as the default information, enabling the classification of users as belonging to one or more of a set of subgroups, and also the integration of the typical characteristics of these subgroups into the individual user profile [132].

Persuasive Technology

Adopting healthy lifestyles is one of the biggest opportunities for preventing chronic diseases. A great interest among researchers is how to explore the use of ICT to change behaviors, a field that has been baptized persuasive technology [133,134]. However, in spite of a wealth of psychological theories on health behavior change, modeling them and embedding them into effective ICT solutions has proved to be difficulty, in great part due to their qualitative nature. Moreover, those theories offer population-based models that do not take into account individual differences. In light of this, there is great promise in combining tracking technologies with statistical techniques to learn an individual's susceptibility to specific persuasive strategies.

Early findings suggest that combining tracking and statistical learning, user's psycho-emotional characteristics and context information, could personalize persuasive technology [135], greatly improving its effectiveness [134]. For instance, where some users might be susceptible to facts and statistics, others will be more easily convinced by emotions and personal stories. By having context aware systems simultaneously, the right moment of messaging can also be selected (eg, when the received message is most necessary and actionable). The field is open for the exploration and exploitation of right-mood right-time right-place feedback that could be more effective than real-time feedback. The goal is to provide evidence of the adherence of people to healthier behaviors using ICT-enabled persuasive services to support safe, secure, seamless monitoring

and persuasive guidance, and personalized assistance from lifestyle improvement to primary, secondary, and tertiary prevention and care [136].

Biomarkers in Cognitive Function Assessment for Wellness

Lifestyle choices greatly affect healthy aging. Still, accurate quantitative diagnostic tests assessing neurodegenerative diseases and cognitive function remain a challenging problem [137]. Currently, diagnosis is mostly conducted by eliminating other possible causes and is usually performed through a combination of psychological tests. However, there has been a significant effort towards the manufacture and market of low-cost consumer brain-computer interface devices [138,139] capable of capturing brain signals and decoding conscious thoughts or emotions, such as movement intentions, facial expressions, excitement, engagement, and frustration. The underlying technology of these wireless devices is based on the ability of the electroencephalogram (EEG) to capture oscillatory brain activity reflecting mean electrical patterns characterizing different brain processes, in terms of cognitive engagement [140-144]. Capturing such activity may be attributed to biomarkers since they reflect the integrity of cognitive pathways. Hence, such information has great potential for use in disease prognosis, in progression monitoring, and even in self-improvement/self-regulation, based on functional connectivity analysis algorithms [143] and neurofeedback approaches [145], respectively. Areas of successful applications include neurodegenerative diseases such as dementia [144], mental disorders such as schizophrenia [146], neurodevelopmental disorders such as autism [147], and even addictive disorders such as alcoholism [148]. Although there is still a long way to go in order to clinically validate these approaches, such technologies are ushering in an era where we will be able to take advantage of the brain's wiring.

Disease Management Using Smart Environment and Personalized Mobile Monitoring Services

New care models incorporating advanced ICTs have the potential to provide service platforms able to improve health care, personalization, inclusion, and empowerment of the individual [149]. Continuous management of diabetic patients can help in achieving effective glucose control and lifestyle changes leading to improved nutrition and healthy levels of physical activity, and to recognize and treat complications [150]. The goal is to empower patients by increasing their ability to self-manage, improve the quality of their life and the overall management of their condition but also to reduce the risk of developing complications and decrease utilization of health care resources [151]. However, there is lack of robust evaluation, including controlled trials on the effectiveness of these powerful new ICT tools [152].

In an Ambient Intelligence environment, various wireless and wired sensor technologies can be integrated, allowing the user to control and interact with the environment. Such innovative systems [152-154] are able to augment surrounding spaces with smart features to allow better lifestyle monitoring and wellness for a user, regardless of geographic location. If the user is inside a hospitalized environment, the system is able to alter the

interfaces to support personalized interaction for the medical personnel. A major challenge related to caring chronic patients is the early detection of exacerbations of the disease. To address this challenge, recent research [155] presents a real-time remote monitoring framework enhanced with semantic technologies (within an Ambient Intelligent environment). It provides personalized, accurate, and fully automated emergency alerting systems that smoothly interact with the personal health professional, regardless of their physical location in order to ensure in-time intervention in case of an emergency.

Smart and Medical Grade Networking for eHealth

Future health informatics for personalized eHealth services rely on innovative technologies for transparent and continuous collection of evidence-based medical information at any time, from anywhere and despite the coverage and availability of communication means. In light of this transformation and change, ICT serves as the catalyst for accelerating the preparedness of all traditional players and to assist all actuators to evolve in envisaged future health care models and systems. Disruption and delay tolerant networking is a novel approach for next-generation eHealth information services where end-to-end homogeneous networking connectivity is not always available [155,156]. For future eHealth and mHealth services, the goal is to provide in-transit persistent information transfer and/or storage allowing uninterruptible services overcoming network instabilities, incompatibilities, or even absence for long periods of time [2].

Personal Health Record Platforms for Digital Patients

Personal health record (PHR) systems are an important, innovative, and constantly evolving area that empowers patients to take a more active role in their own health and make informed decisions. One of the most promising aspects of PHRs is to improve health care delivery and reduce costs. The primary goal of a PHR system is to provide the patient with the ability to maintain and manage their PHR, a "systematic collection of information about an individual's health and health care, stored in electronic format" [129-131,157]. PHRs provide a complete summary of patients' health history, enhance accurate clinical diagnosis, and empower patients to manage their own health [131]. The interconnection of data sources is an important aspect for modern PHRs. The collection of heterogeneous health parameters, for accessing, sharing, and analyzing long-term multilevel health data, including clinical, genetic, sensor, human behavior, and activity, enables clinical analysis, prediction, and prevention for the individual citizen. Triggering intervention on detection of conditions that may lead to health deterioration for preventive care becomes possible. Recently, there have also been development efforts that aim to implement a useful, effective, and intelligent PHR framework that will satisfy the variety of health environment needs and foster an optimal user experience [107,157].

Drawbacks and Pitfalls

Technological innovations to foster personalized healthy living and wellness could dramatically improve our ability to sense, monitor, and manage our health status and contribute to a change of paradigm in health care. Yet, misuse of technology and, above

all, personal data generated through them, may lead to substantial deleterious effects. Below, we discuss important drawbacks and pitfalls that must be addressed for the use of services supported in such technologies to be perceived as secure, safe, and trustable and to be adopted at a greater level than today.

Security and Trust

The objective of a secure system is to protect sensitive health information from unauthorized access, manipulation, and misuse. The protection goals that ascribe the requirements of a secure system are defined by the following objectives: authentication, authorization, confidentiality, integrity, accountability, and non-repudiation. It is essential that intelligent pervasive health care solutions are developed and correctly integrated to assist health care professionals in delivering high levels of patient care [158]. It is equally important that these solutions are used to empower patients and relatives for self-care or wellness management and to provide seamless, trustable access to health care services. One of the major challenges to be addressed is how to ensure security and privacy. Adapt-lite [159] and Hide-n-sense [97] are examples of how security mechanisms can be applied in mHealth services.

The face of health care is changing as new technologies are being incorporated into the existing infrastructure. The population of new technologies could possible dramatically improve our ability to detect, monitor, and address lifestyles and wellness but also could cause substantial deleterious effects. Adoption of technological innovation to support personalized healthy lifestyles and well-being depends on the extent that public concerns about privacy, confidentiality, and security of online data are addressed. Today there is a rapidly growing market of online apps and social media tools for health, with little focus on the issues of ownership and protection of data [160].

A huge volume of data is generated and is expected to increase by the use of future services either intentionally by users or automatically by networked devices, such as information appliances and sensors. Product developers and policy makers will need to proactively balance public concerns about privacy protection with the information-sharing needs of some business models and public health programs. Policies, regulatory and otherwise, will need to keep pace with technologic innovation.

A cloud infrastructure, for example, is susceptible to many privacy and security attacks [161]. As a result, many hospitals, biomedical research groups and health care organizations are reluctant to adopt this technology as a privacy breach with respect to the patient information managed under their jurisdiction could be devastating, especially in terms of cost [162]. This is highlighted in the work of Seong Han et al [163], who report that the biggest obstacle hindering the progress and the wide adoption of the cloud are privacy and security issues associated with it. This is further evidenced from a survey carried out by IDC Enterprise Panel [164], where most users pointed out security as the top challenge. Nevertheless, significant research still needs to be done to ensure that only selected and trusted resources are used, for example, a

trust-based security framework as proposed by van't Noordende et al [165].

In order to foster trust, users must receive clear and objective information about the benefits and risks associated with the use of the new systems and services. Since sensitive health and lifestyle data are going to be processed, measures must be taken in order to protect the data, stored and in transit, against unauthorized disclosure or access, accidental or unlawful destruction, or accidental loss or alteration according to Article 17 (1) of the Data Protection Directive (Directive 95/46/EC). When one considers the requirements imposed by standards or other regulations, such as the requirements for the fair and lawful processing of data established by the European Data Protection Directive [166] or the requirements of HIPAA (Health Insurance Portability and Accountability Act) [167] in the United States, it becomes evident that it is crucial for health-related data to be kept confidential from anyone unless authorized by the patient or some emergency regulations. Moreover, establishing a legal framework of future services is required in order to ensure lawful and fair data processing for personal medical and clinical data. On January 25, 2012, the EU Commission published its draft General Data Protection Regulation 2012/0011(COD). The Regulation's stated intention is to build a stronger and more coherent data protection framework in the EU that will resolve current legal uncertainties, put individuals in control of their own data, and bring greater legal and practical certainty for organizations that are subject to the legislation. The EU Commission has indicated that it aims to have in place a revised legislative framework by 2016.

Quality and Effectiveness

Many of these innovations will be complex and would need to be networked to and operate in conjunction with other existing services and applications. Ensuring interoperability is fundamental in order to minimize the risk of disruption when integrating heterogeneous services. This can be achieved by using quality and evaluation processes throughout the development life cycle and stable architectural description plans [168]. In order to foster effectiveness, we must redesign the methods and context for performing real-life trials to ensure the evaluation of these technologies and stretch the limits of existing methods of creating and applying robust interventions. The ultimate goal is to identify and answer the health and social impact of the intervention at the population level. The challenge will be to develop consensus methods and metrics around this fundamental question. Today, many services widely deployed have been validated using methodologies inherited from traditional health interventions and many times do not have empirical evidence of benefit [169].

Discussion

Principle Considerations

Advances in the theory of the brain, neuroscience and the mechanisms of behavior modification [137,143] and the power of positive psychology offer an opportunity to health care providers and managers, as well as citizens for real empowerment to change the biggest and yet modifiable factors that contribute to the top chronic diseases. The economic impact

of these chronic diseases and the "ongoing economic uncertainty brings into sharp focus the fact that current health care models are financially unsustainable" [2]. Moreover, the traditional health care system alone has not been able to accomplish the shift toward healthier lifestyles and the expected reduction in chronic diseases and their associated economic burden.

ICT offers a promising new approach toward wellness throughout the lifespan, by fostering and supporting the active and meaningful participation of individuals in their health management. At the same time, it can maximize the information flow with clinicians and open new horizons for individualized medicine approaches based on multilevel personal data.

In order to realize this paradigm shift towards citizen empowerment and engagement in health and well-being production, it is clear that a new generation of ubiquitous and convergent network and service infrastructures is required. Their function will be to sustain the construction and the deployment of highly personalized, scalable, flexible, manageable, context-aware, dependable, and secure services incorporating resources in a holistic seamless ecosystem [150]. These infrastructures need to support an Internet of things, dynamically combining devices, communication and delivery systems and services. Virtualization of resources remains an important research driver, enabling the delivery of services independently from the underlying platform [2,170]. The focus is on the use of pervasive mobile technologies so that scientists and researchers can easily design, share, and execute simulations.

Enormous challenges remain and multidisciplinary approaches are required to engineer an information infrastructure that ensures privacy and supports intelligent and personalized user-computer interaction, innovative conversional agents and assistants with emphasis on a citizen's engagement and empowerment. Computer-delivered interventions can lead to better behavioral health outcomes, post-intervention improvements in health-related knowledge, attitudes, and intentions as well as positive changes in health behaviors such as dietary intake, tobacco use, substance use, safer sexual behavior, binge/purging behaviors, and general health maintenance [171].

The primary challenge for relevant future innovations is to utilize and modify existing technology with an emphasis on mobile communication technology and devices to (1) create new models of impact on chronic disease development able to empower individuals to adopt and sustain changes towards healthy choices (eg, nutritional habits, physical activity, stress management, sleep, anxiety, depression, meaningful social interaction, and networking) and to inform decisions makers and funding bodies on the best available choices; (2) create positive and personally motivating marketing strategies for individuals to become engaged and motivated to improve their own health on the following pillars: sensing and assessing individualized health behaviors and status, providing feedback tailored to psychosocial profiles, incorporating end users in the development of the technology, creating strong relationships for ownership and empowerment, and connecting the individual to community-based support/capacity structures to boost, encourage, sustain, and integrate their own self-regulating

efforts; (3) use and improve available health supporting services for sensing and early detection of risky health behaviors; (4) develop an ecosystem of community-based capacity to support individuals who choose to be coproducers of healthy lifestyles in all levels of health care delivery (primary, secondary, tertiary); (5) provide a "proof of concept" of this model of integrated person-centered healthier lifestyle and well-being through ICT; and (6) create appropriate business models of community-integrated elements for health coproduction across cultural milieu, able to foster the scaling up of results and to bring social and technological innovation to the market.

In achieving such a paradigmatic shift, it is critical to provide smart ubiquitous services and systems advancing the current state-of-the-art in a number of domains relevant for risk factors that may influence directly or indirectly an individual's wellness and well-being and lifestyle behaviors: (1) smart and ubiquitous sensor fusion and Internet of things services to allow health-related information to be aggregated and transmitted for remote monitoring and response, to support personalized/individualized multilevel patient/citizen-centered health care services; (2) assistant/coaching services for wellness management enabling citizens to manage their health and well-being; (3) intelligent decision-making support services through the combination of understanding with the ability to influence decisions by communicating subjective values, helping to select interventions, assessing outcomes, and providing feedback to the user; (3) novel mechanisms and services to support self-efficacy confidence that the individual can perform a given behavior, including decision making and a belief in their ability to change the situation; (4) ubiquitous just-in time support services with mechanisms that incorporate known principles of health behavior change, based on the user's biopsychosocial profiling with representations of motivation cues and techniques for overcoming barriers to change; (5) ecosystems and personalized health promoting services managing multilevel information and medical data from multidisciplinary domains, including the health and social care systems; (6) service infrastructures able to support unified data access, management, presentation, sharing, and security, based on specific user requirements and engagement; (7) linkage services, able to connect government agencies and authorities to allow official public health structures to use the coming innovation for preventing diseases and promoting health through organized efforts and informed societal choices of public and private organizations, communities, and individuals; (8) train for the health and social care workforce in the utility of ICT; and (9) programs able to generate large population-based evidence on the efficacy and cost effectiveness of this paradigm shift.

Conclusion

This paper focuses on innovative and meaningful ways to empower and engage individual citizens into sharing knowledge and awareness and becoming coproducers of their health and wellness through the adoption of eHealth and mHealth technology. This personal choice to being coproducer of one's own health and wellness will have a ripple effect in both the way health is produced and maintained as well as on the kind of models and enterprises that can sprout in the communities of individual users.

Despite the innovative approach and expected results, such a paradigmatic shift cannot happen by overriding existing wellness and chronic disease management practices and services. Change needs to happen with and within current arrangements. In order to support and strengthen such movement, a number of trends would need to be supported by future research: (1) Evidence depends on the implementation of large-scale solutions based on innovative services, by means of a patient-centered approach that must also address the issue of comorbidities and their management; (2) Outcomes must be validated, strengthening the evidence for chronic disease management, through robust health technology assessment for effectiveness, quality, and continuity of care, personalization, cost-efficiency, satisfaction of the users, and transferability; (3) The rollout of services should be preceded by the development of guidelines to identify profiles of patients who may benefit from the provision of such services (condition, age, severity of the condition, comorbidity, socioeconomic status, and any other relevant factors); (4) Conditions for building on top of existing solutions and reuse, when possible, established and scientifically validated methodologies must be studied, and public national and regional authorities that act today as eHealth service providers must be involved; (5) Scale-up will depend on the dissemination and exploitation of good practices and coaching of early adopters and followers on how to effectively connect to the existing traditional health care system; (6) Assuring security and privacy in digital environments through stronger and transparent mechanisms, as predecessors of safety and trust, is of fundamental importance; providers need to realize that trust is a factor perceived by the user based on a dynamic process and it must be cared over time; and (7) The complex, changing needs of a mobile society will increasingly push for an agenda where all the above aspects will meet under cross-border deployment of health and wellness management services, and issues of interoperability, quality, and care standards will regain importance.

The challenge today is to facilitate the adoption and use of ICT tools that go beyond the current state-of-the-art, as described in this paper, not only by the engaged young population but also by the unengaged older people who are in urgent need of behavior change towards healthier life habits. We need health data to flow back to its owners so they can manage their own health and improve their own health metrics. The evidence so far suggests that these services need to be easily accessible, enjoyable, connected to people's everyday life activities, and adapted to the end-user's physiological and age-related changes. It is also important to educate the medical and social systems on the use and benefits of these technologies for achieving the desired health outcomes. By introducing such innovations within the health care system, we can promote community capacity and social engagement through better and healthier behavior and lifestyle of individuals.

Authors' Contributions

EGS and SS have contributed to conception and design, acquisition of data, analysis and interpretation of data, drafting or revising the manuscript critically for important intellectual content, and final approval of the version to be published. MT, VS, KM, JHJ, HdJ, and AT have contributed, by order of names in the list of authors, to the analysis and interpretation of data and revising the manuscript. CT has contributed to the conception and design, acquisition of data, analysis and interpretation of data, and drafting or revising the manuscript critically for important intellectual content, especially from the medical point of view.

Conflicts of Interest

None declared.

References

1. World Health Organization. Global Health Observatory Data Repository. 2015 Jan. URL: http://www.who.int/mediacentre/factsheets/fs355/en/ [accessed 2016-05-11] [WebCite Cache ID 6hQS92V9G]
2. eHealth Task Force Report. Redesigning health in Europe for 2020. Belgium: Luxembourg: Publications Office of the European Union, 2012; 2012.
3. Kondylakis H, Spanakis EG, Sfakianakis S, Sakkalis V, Tsiknakis M, Marias K, et al. Digital patient: Personalized and translational data management through the MyHealthAvatar EU project. Conf Proc IEEE Eng Med Biol Soc 2015 Aug:1397-1400. [doi: 10.1109/EMBC.2015.7318630] [Medline: 26736530]
4. Samb B, Desai N, Nishtar S, Mendis S, Bekedam H, Wright A, et al. Prevention and management of chronic disease: a litmus test for health-systems strengthening in low-income and middle-income countries. Lancet 2010 Nov 20;376(9754):1785-1797. [doi: 10.1016/S0140-6736(10)61353-0] [Medline: 21074253]

XSL•FO
RenderX

5.   Lawrence OG. Healthy Living Needs Global Governance. Nature 2014 Jul 10;511:147-150 [FREE Full text] [doi: 10.1016/S0140-6736(10)61353-0]

6.   Ratzan SC. Chronic disease: a risk we need to score. J Health Commun 2010 Jul;15(5):467-469. [doi: 10.1080/10810730.2010.495008] [Medline: 20677052]

7.   Marteau TM, Ogilvie D, Roland M, Suhrcke M, Kelly MP. Judging nudging: can nudging improve population health? BMJ 2011;342:d228. [Medline: 21266441]

8.   Dunton GF, Atienza AA, Castro CM, King AC. Using ecological momentary assessment to examine antecedents and correlates of physical activity bouts in adults age 50+ years: a pilot study. Ann Behav Med 2009 Dec;38(3):249-255. [doi: 10.1007/s12160-009-9141-4] [Medline: 20052568]

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

10.   Büschel I, Mehdi R, Cammilleri A, Marzouki Y, Elger B. Protecting human health and security in digital Europe: how to deal with the "privacy paradox"? Sci Eng Ethics 2014 Sep;20(3):639-658. [doi: 10.1007/s11948-013-9511-y] [Medline: 24446151]

11.   van Gemert-Pijnen J, Nijland N, van LM, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]

12.   Choudhury SR, Barman A. Technology and Well-Being- An Evocative Essay. Postmodern Openings 2014 Jun;5(2):15-37.

13.   Foster L, Keller PC, McKee B, Ostry A. The British Columbia Atlas of Wellness (2nd Edition). Victoria, Canada: Western Geographical Press; 2011.

14.   Adams T, Bezner J, Steinhardt M. The conceptualization and measurement of perceived wellness: integrating balance across and within dimensions. Am J Health Promot 1997;11(3):208-218. [Medline: 10165100]

15.   Anspaugh DJ, Hamrick MH, Rosato FD. Wellness: Concepts and Applications (8th Edition). NY, USA: McGraw-Hill Humanities/Social Sciences/Languages; 2011.

16.   Crose R, Nicholas DR, Gobble DC, Frank B. Gender and Wellness: A Multidimensional Systems Model for Counseling. Journal of Counseling & Development 1992 ;71(2):149-156 [FREE Full text] [doi: 10.1002/j.1556-6676.1992.tb02190.x]

17.   Durlak J. Health promotion as a strategy in primary prevention. In: Cichetti D, Rappaport J, Sandler I, Weissberg R, editors. The promotion of wellness in children and adolescents. Washington, DC: Child Welfare League Association Press; 2000:221-241.

18.   Hales D. An Invitation to health (11th ed). Belmont, CA: Thomson & Wadsworth; 2005.

19.   Helliwell JF. Well-Being, Social Capital and Public Policy: What's New? Economic Journal 2006 Mar;116(510):C34-C45. [doi: 10.1111/j.1468-0297.2006.01074.x]

20.   Hettler B. Wellness promotion on a university campus. Fam Community Health 1980 May;3(1):77-95. [Medline: 10246133]

21.   Leafgren F. Being a man can be hazardous to your health: Life-styles issues. In: Moore D, Leafgren F, editors. Problem solving strategies and interventions for men in conflict. USA: American Association for Counseling and Development; 1990:265-311.

22.   Renger RF, Midyett SJ, Soto Mas FG, Erin TD, McDermott HM, Papenfuss RL, et al. Optimal Living Profile: An Inventory to Assess Health and Wellness. Am J Health Behav 2000 Nov 01;24(6):403-412. [doi: 10.5993/AJHB.24.6.1]

23.   Ryan RM, Deci EL. On happiness and human potentials: a review of research on hedonic and eudaimonic well-being. Annu Rev Psychol 2001;52:141-166. [doi: 10.1146/annurev.psych.52.1.141] [Medline: 11148302]

24.   Fitzgerald N, Spaccarotella K. Barriers to a healthy lifestyle: From individuals to public policy - an ecological perspective. Journal of Extension 2009 Feb;47(1):1-8. [Medline: 11148302]

25.   Bevan AL. Creating communicative spaces in an action research study. Nurse Res 2013 Nov;21(2):14-17. [doi: 10.7748/nr2013.11.21.2.14.e347] [Medline: 24171632]

26.   Duff C. Networks, resources and agencies: on the character and production of enabling places. Health Place 2011 Jan;17(1):149-156. [doi: 10.1016/j.healthplace.2010.09.012] [Medline: 20926330]

27.   Ryff CD, Singer BH. Best news yet on the six-factor model of well-being. Social Science Research 2006 Dec;35(4):1103-1119. [doi: 10.1016/j.ssresearch.2006.01.002]

28.   Lauren JR. Wellness: A review of theory and measurement for counselors. Journal of Counseling & Development 2009;87(3):216-226. [doi: 10.1002/j.1556-6678.2009.tb00570.x]

29.   Gibbs WW. Medicine gets up close and personal. Nature 2014 Feb 13;506(7487):144-145. [doi: 10.1038/506144a] [Medline: 24522582]

30.   MacLeod J, Nelson G. Programs for the promotion of family wellness and the prevention of child maltreatment: a meta-analytic review. Child Abuse Negl 2000 Sep;24(9):1127-1149. [Medline: 11057701]

31.   Hallal PC, Bauman AE, Heath GW, Kohl HW, Lee IM, Pratt M. Physical activity: more of the same is not enough. Lancet 2012 Jul 21;380(9838):190-191. [doi: 10.1016/S0140-6736(12)61027-7] [Medline: 22818932]

XSL·FO

RenderX

32.   Beck T, Gollapudi S, Brunak S, Graf N, Lemke HU, Dash D, et al. Knowledge engineering for health: a new discipline required to bridge the "ICT gap" between research and healthcare. Hum Mutat 2012 May;33(5):797-802. [doi: 10.1002/humu.22066] [Medline: 22392843]

33.   Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. J Med Internet Res 2014;16(8):e192 [FREE Full text] [doi: 10.2196/jmir.3469] [Medline: 25131661]

34.   Wayne N, Ritvo P. Smartphone-enabled health coach intervention for people with diabetes from a modest socioeconomic strata community: single-arm longitudinal feasibility study. J Med Internet Res 2014;16(6):e149 [FREE Full text] [doi: 10.2196/jmir.3180] [Medline: 24907918]

35.   Chiarugi F, Trypakis D, Spanakis EG. Problems and Solutions for Storing and Sharing Data from Medical Devices in eHealth Applications. 2004 Presented at: 2nd OpenECG Workshop; April 1-3, 2004; Berlin, Germany.

36.   Mohr DC, Schueller SM, Montague E, Burns MN, Rashidi P. The behavioral intervention technology model: an integrated conceptual and technological framework for eHealth and mHealth interventions. J Med Internet Res 2014;16(6):e146 [FREE Full text] [doi: 10.2196/jmir.3077] [Medline: 24905070]

37.   Hwang S, Yu D. GPS localization improvement of smartphones using built-in sensors. International Journal of Smart Home 2012;6(3):1-8.

38.   Thomas JG, Wing RR. Health-e-call, a smartphone-assisted behavioral obesity treatment: pilot study. JMIR Mhealth Uhealth 2013;1(1):e3 [FREE Full text] [doi: 10.2196/mhealth.2164] [Medline: 25100672]

39.   Chiarugi F, Spanakis EG, Lees PJ, Chronaki CE, Tsiknakis M, Orphanoudakis SC. ECG in your hands: A multi-vendor ECG viewer for personal digital assistants. 2003 Presented at: Computers in Cardiology; Sept. 21-24, 2003; Thessaloniki, Chalkidiki, Greece p. 359-362.

40.   Pagoto S, Schneider K, Jojic M, DeBiasse M, Mann D. Evidence-based strategies in weight-loss mobile apps. Am J Prev Med 2013 Nov;45(5):576-582. [doi: 10.1016/j.amepre.2013.04.025] [Medline: 24139770]

41.   Lane ND, Mohammod M, Lin M, Yang X, Lu H, Ali S, et al. Bewell: A smartphone application to monitor, model and promote wellbeing. 2011 Presented at: 5th International ICST Conference on Pervasive Computing Technologies for Healthcare; May 23-26, 2011; Dublin, Ireland p. 23-26.

42.   Common Sense for Developers. URL: http://developer.sense-os.nl/ [accessed 2015-06-09] [WebCite Cache ID 6Z9xOmpm2]

43.   Funf Open Sensing Framework. URL: http://www.funf.org/ [accessed 2015-06-09] [WebCite Cache ID 6Z9xRZHZh]

44.   Ubifit. URL: http://dub.washington.edu/projects/ubifit [accessed 2015-06-09] [WebCite Cache ID 6Z9xTeSnW]

45.   Lane ND, Xu Y, Lu H, Choudhury T, Campbell AT, Zhao F. Enabling large-scale human activity inference on smartphones using community similarity networks. 2011 Jan 9 Presented at: 13th International Conference on Ubiquitous Computing, ACM; September 17-21, 2011; New York, NY p. 355-364.

46.   Aram S, Troiano A, Pasero E. Environment sensing using smartphone. 2012 Feb 7 Presented at: IEEE Sensors Applications Symposium (SAS); Feb. 7-9, 2012; Brescia p. 1-4.

47.   Aleksandar M, Osmani V, Popleteev A, Mayora-Ibarra O. Smart Phone Sensing to Examine Effects of Social Interactions and Non-sedentary Work Time on Mood Changes. In: Lecture Notes in Computer Science. Karlsruhe, Germany: Springer-Verlag; 2011 Presented at: Modeling and Using Context - 7th International and Interdisciplinary Conference, CONTEXT 2011; September 26-30, 2011; Berlin p. 200-213.

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

49.   Samsung. S Health. URL: http://content.samsung.com/uk/contents/aboutn/sHealthIntro.do [WebCite Cache ID 6Z9xn8SuX]

50.   Apple Health. URL: https://www.apple.com/ios/whats-new/health/ [accessed 2015-06-09] [WebCite Cache ID 6Z9xrImd1]

51.   Motoactv. URL: https://motoactv.com/ [accessed 2015-06-09] [WebCite Cache ID 6Z9xtt6Eh]

52.   Samsung Gear. URL: http://www.samsung.com/global/microsite/gear/index.html [accessed 2015-06-09] [WebCite Cache ID 6Z9xy5ljK]

53.   Apple Watch. URL: http://www.apple.com/watch/ [accessed 2015-06-09] [WebCite Cache ID 6Z9y07s6c]

54.   Moto360. URL: https://moto360.motorola.com/ [accessed 2015-06-09] [WebCite Cache ID 6Z9y2IC7a]

55.   Modave F, Bian J, Leavitt T, Bromwell J, Harris IC, Vincent H. Low Quality of Free Coaching Apps With Respect to the American College of Sports Medicine Guidelines: A Review of Current Mobile Apps. JMIR Mhealth Uhealth 2015;3(3):e77 [FREE Full text] [doi: 10.2196/mhealth.4669] [Medline: 26209109]

56.   Shaw RJ, Steinberg DM, Bonnet J, Modarai F, George A, Cunningham T, et al. Mobile health devices: will patients actually use them? J Am Med Inform Assoc 2016 Jan 17:ocv186. [doi: 10.1093/jamia/ocv186]

57.   Illiger K, Hupka M, von JU, Wichelhaus D, Albrecht U. Mobile technologies: expectancy, usage, and acceptance of clinical staff and patients at a university medical center. JMIR Mhealth Uhealth 2014;2(4):e42 [FREE Full text] [doi: 10.2196/mhealth.3799] [Medline: 25338094]

58.   Brdiczka O, Crowley JL, Reignier P. Learning Situation Models for Providing Context-Aware Services. 2007 Presented at: HCI 2007; July 22-27, 2007; Beijing, China p. 23-32.

XSL•FO

RenderX

59. Alshurafa N, Xu W, Liu JJ, Huang M, Mortazavi B, Roberts CK, et al. Designing a robust activity recognition framework for health and exergaming using wearable sensors. IEEE J Biomed Health Inform 2014 Sep;18(5):1636-1646. [doi: 10.1109/JBHI.2013.2287504] [Medline: 24235280]

60. Fatima I, Fahim M, Lee Y, Lee S. A unified framework for activity recognition-based behavior analysis and action prediction in smart homes. Sensors (Basel) 2013;13(2):2682-2699 [FREE Full text] [doi: 10.3390/s130202682] [Medline: 23435057]

61. Broekens J, Bosse T, Marsella SC. Challenges in computational modeling of affective processes. IEEE Trans. Affective Computing 2013;4:242-245.

62. Janssen J, Tacken P, de Vries J, van den Broek E, Westerink J, Haselager P, et al. Machines Outperform Laypersons in Recognizing Emotions Elicited by Autobiographical Recollection. Human–Computer Interaction 2013 Nov 02;28(6):479-517. [doi: 10.1080/07370024.2012.755421]

63. Kaptein M, De Ruyter B, Markopoulos P, Aarts E. Adaptive Persuasive Systems. ACM Trans Interact Intell Syst 2012 Jun 01;2(2):1-25. [doi: 10.1145/2209310.2209313]

64. Vinciarelli A, Pantic M, Bourlard H. Social signal processing: Survey of an emerging domain. Image and Vision Computing 2009 Nov;27(12):1743-1759. [doi: 10.1016/j.imavis.2008.11.007]

65. Janssen J. A three-component framework for empathic technologies to augment human interaction. J Multimodal User Interfaces 2012 Apr 14;6(3-4):143-161. [doi: 10.1007/s12193-012-0097-5]

66. Ferreira F, Almeida N, Rosa A, Oliveira A, Casimiro J, Silva S, et al. Elderly Centered Design for Interaction – The Case of the S4S Medication Assistant. In: Procedia Computer Science. 2014 Presented at: 5th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion, DSAI 2013; November 13-15, 2013; Vigo, Spain p. 398-408. [doi: 10.1016/j.procs.2014.02.044]

67. Allen J, Ferguson G, Blaylock N, Byron D, Chambers N, Dzikovska M, et al. Chester: towards a personal medication advisor. J Biomed Inform 2006 Oct;39(5):500-513 [FREE Full text] [doi: 10.1016/j.jbi.2006.02.004] [Medline: 16545620]

68. Allen J, Ferguson G, Stent A. An Architecture For More Realistic Conversational Systems. In: IUI 2001. 2001 Presented at: Conference on Intelligent User Interfaces; 2001; Santa Fe, NM p. 1-8.

69. Ferguson G, Allen J, Galescu L, Quinn J, Swift M. CARDIAC: An intelligent conversational assistant for chronic heart failure, Patient Heath Monitoring. 2009 Presented at: AAAI Fall Symposium on Virtual Healthcare Interaction; November 5–7, 2009; Arlington, Virginia.

70. Foster ME. COMIC project deliverable 6.1. 2002. State of the art review: Multimodal fission URL: http://groups.inf.ed.ac.uk/comic/documents/deliverables/Del6-1.pdf [accessed 2016-05-11] [WebCite Cache ID 6hQzAiizq]

71. Rousseau C, Bellik Y, Vernier F. WWHT: Un modèle conceptuel pour la présentation multimodale d'information. In: IHM 2005. 2005 Presented at: 17th Conference on l'Interaction Homme-Machine; 2005; New York, NY p. 59-66. [doi: 10.1145/1148550.1148558]

72. Chawla NV, Davis DA. Bringing big data to personalized healthcare: a patient-centered framework. J Gen Intern Med 2013 Sep;28 Suppl 3:S660-S665 [FREE Full text] [doi: 10.1007/s11606-013-2455-8] [Medline: 23797912]

73. Rosenthal A, Mork P, Li MH, Stanford J, Koester D, Reynolds P. Cloud computing: a new business paradigm for biomedical information sharing. J Biomed Inform 2010 Apr;43(2):342-353 [FREE Full text] [doi: 10.1016/j.jbi.2009.08.014] [Medline: 19715773]

74. Rubenstein K. Cloud computing in life sciences R&D. Cambridge: Healthtech Institute; 2010 Apr. URL: http://www.informationweek.com/systems-integration/why-it-needs-to-push-data-sharing-efforts/d/d-id/1090052? [accessed 2015-06-09] [WebCite Cache ID 6Z9yQ0FOm]

75. Thilakanathan D, Chen S, Nepal S, Calvo R, Alem L. A platform for secure monitoring and sharing of generic health data in the Cloud. Future Generation Computer Systems 2014 Jun;35:102-113. [doi: 10.1016/j.future.2013.09.011]

76. Basdekis I, Sakkalis V, Stephanidis C. Towards an Accessible Personal Health Record. In: Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering (LNICST).: Springer; 2012 Presented at: MobiHealth 2011; October 5-7, 2011; Kos Island, Greece p. 61-68.

77. Maniadi E, Kondylakis H, Spanakis EG, Tsiknakis M, Marias K, Dong F. Designing a digital patient avatar in the context of the MyHealthAvatar project initiative. 2013 Presented at: IEEE 13th International Conference on BioInformatics and BioEngineering (BIBE); Nov. 10-13, 2013; Chania, GR. [doi: 10.1109/BIBE.2013.6701560]

78. Giniat EJ. Cloud computing: innovating the business of health care. Healthc Financ Manage 2011 May;65(5):130-131. [Medline: 21634276]

79. Healey M. Information Week. 2010. Why IT needs to push data sharing efforts URL: http://www.itone.lu/article/why-it-needs-push-data-sharing-efforts [accessed 2016-05-11] [WebCite Cache ID 6hR01RKA5]

80. Wu R, Ah GJ, Hu H. Secure sharing of electronic health records in clouds. In: CollaborateCom 2012. 2012 Presented at: 8th International Conference on Collaborative Computing: Networking, Applications and Worksharing; Oct. 14-17, 2012; Pittsburgh, PA p. 711-718.

81. Sarathy R, Muralidhar K. Secure and useful data sharing. Decision Support Systems 2006 Oct;42(1):204-220. [doi: 10.1016/j.dss.2004.10.013]

82. Butler D. Data sharing threatens privacy. Nature 2007 Oct 11;449(7163):644-645. [doi: 10.1038/449644a] [Medline: 17968386]

83. Feldman L, Patel D, Ortmann L, Robinson K, Popovic T. Educating for the future: another important benefit of data sharing. Lancet 2012 May 19;379(9829):1877-1878. [doi: 10.1016/S0140-6736(12)60809-5] [Medline: 22608336]

84. Rocha F, Abreu S, Correia M. The Final Frontier: Confidentiality and Privacy in the Cloud. Computer 2011 Sep;44(9):44-50. [doi: 10.1109/MC.2011.223]

85. Duncan A, Creese S, Goldsmith M. Insider Attacks in Cloud Computing. 2012 Presented at: IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications; June 25-27, 2012; Liverpool, UK.

86. National Institute of Standards and Technology (NIST). Special Publication 800-144.: US Department of Commerce Guidelines on security and privacy in public cloud computing URL: http://csrc.nist.gov/publications/nistpubs/800-144/SP800-144.pdf [accessed 2015-06-09] [WebCite Cache ID 6Z9yZ8kh2]

87. Li M, Yu S, Zheng Y, Ren K, Lou W. Scalable and Secure Sharing of Personal Health Records in Cloud Computing Using Attribute-Based Encryption. IEEE Trans. Parallel Distrib. Syst 2013 Jan;24(1):131-143. [doi: 10.1109/TPDS.2012.97]

88. Tu SS, Niu SZ, Hui L, Xiao-Ming Y, Li MJ. Fine grained access control and revocation for sharing data on clouds. In: IPDPSW 2012. 2012 Presented at: IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum; May 21-25, 2012; Shanghai. [doi: 10.1109/IPDPSW.2012.265]

89. Continua Alliance. URL: http://www.continuaalliance.org/ [accessed 2016-05-18] [WebCite Cache ID 6hbJbcSMK]

90. Bluetooth USB dongle. URL: http://www.stollmann.de/en/modules/bluetooth-products/bluehdp-usb.html [accessed 2015-06-09] [WebCite Cache ID 6Z9z6wgLa]

91. Ant+. URL: http://www.thisisant.com/ [accessed 2015-06-09] [WebCite Cache ID 6Z9zBNPM3]

92. Bierhoff I, Rijnen W, Goossen B. Easy, familiar, enjoyable and empowering netclarity services. In: AAL Forum 2011. 2011 Presented at: Partnerships for Social Innovation in Europe; September 26-28, 2011; Lecce, Italy p. 301-302.

93. Omar WM, Taleb-Bendiab A. E-health support services based on service-oriented architecture. 2006 Presented at: 2006 IEEE International Conference on Services Computing (SCC'06); Sept. 18-22, 2006; Chicago, IL p. 135-142.

94. Hristoskova A, Sakkalis V, Zacharioudakis G, Tsiknakis M, De Turck TF. Ontology-driven monitoring of patient's vital signs enabling personalized medical detection and alert. Sensors (Basel) 2013;14(1):1598-1628 [FREE Full text] [doi: 10.3390/s140101598] [Medline: 24445411]

95. Benharref A, Serhani MA. Novel cloud and SOA-based framework for e-health monitoring using wireless biosensors. IEEE J Biomed Health Inform 2014 Jan;18(1):46-55. [doi: 10.1109/JBHI.2013.2262659] [Medline: 24403403]

96. de Deugd S, Carroll R, Kelly K, Millett B, Ricker J. SODA: Service Oriented Device Architecture. IEEE Pervasive Comput 2006 Jul;5(3):94-96, c3. [doi: 10.1109/MPRV.2006.59]

97. Mare S, Sorber J, Shin M, Cornelius C, Kotz D. Hide-n-Sense: Privacy-aware secure mHealth sensing. 2011. Dartmouth Computer Science Technical Report URL: http://sharps.org/wp-content/uploads/MARE-TR.pdf [accessed 2016-05-18] [WebCite Cache ID 6hbdqMTzX]

98. GoPubMed. URL: http://www.gopubmed.org/web/gopubmed/ [accessed 2015-06-09] [WebCite Cache ID 6Z9zGUzTc]

99. Geneontology Consortium. URL: http://www.geneontology.org [WebCite Cache ID 6Z9zJKU5Y]

100. Medical Subject Headings. URL: http://www.nlm.nih.gov/mesh [WebCite Cache ID 6Z9zMdLGp]

101. Kupershmidt I, Su QJ, Grewal A, Sundaresh S, Halperin I, Flynn J, et al. Ontology-based meta-analysis of global collections of high-throughput public data. PLoS One 2010;5(9):- [FREE Full text] [doi: 10.1371/journal.pone.0013066] [Medline: 20927376]

102. Teixeira A, Rodrigues M, Ferreira L. Online Health Information Semantic Search and Exploration: Reporting on two prototypes for performing information extraction on both a hospital Intranet and on the World Wide Web. In: Neustein A, editor. Speech Technology and Text Mining in Medicine and Healthcare. Berlin, Germany: De Gruyter; 2014.

103. Wroe C, Stevens R, Goble C, Roberts A, Greenwood M. A suite of DAML+OIL ontologies to describe bioinformatics web services and data. Int J Coop Info Syst 2003 Jun;12(02):197-224. [doi: 10.1142/S0218843003000711]

104. Spring B, Ockene JK, Gidding SS, Mozaffarian D, Moore S, Rosal MC, American Heart Association Behavior Change Committee of the Council on Epidemiology and Prevention, Council on Lifestyle and Cardiometabolic Health, Council for High Blood Pressure Research, Council on Cardiovascular and Stroke Nursing. Better population health through behavior change in adults: a call to action. Circulation 2013 Nov 5;128(19):2169-2176 [FREE Full text] [doi: 10.1161/01.cir.0000435173.25936.e1] [Medline: 24100544]

105. Cubo J, Nieto A, Pimentel E. A cloud-based Internet of Things platform for ambient assisted living. Sensors (Basel) 2014;14(8):14070-14105 [FREE Full text] [doi: 10.3390/s140814070] [Medline: 25093343]

106. Hafen E, Kossmann D, Brand A. Health data cooperatives - citizen empowerment. Methods Inf Med 2014;53(2):82-86. [doi: 10.3414/ME13-02-0051] [Medline: 24514946]

107. Spanakis EG, Kafetzopoulos D, Yang P, Marias K, Deng Z, Tsiknakis M, et al. MyHealthAvatar: personalized and empowerment health services through Internet of Things technologies. 2014 Presented at: EAI 4th International Conference on Wireless Mobile Communication and Healthcare (Mobihealth); November, 2014; Athens, Greece p. 3-5.

108. Honka A, Kaipainen K, Hietala H, Saranummi N. Rethinking health: ICT-enabled services to empower people to manage their health. IEEE Rev Biomed Eng 2011;4:119-139. [doi: 10.1109/RBME.2011.2174217] [Medline: 22273795]

XSL•FO
RenderX

109. Chang KH, Liu SY, Chu HH, Hsu JY, Chen C, Lin TY, et al. The diet-aware dining table: observing dietary behaviors over a tabletop surface. In: PERVASIVE 2006, Proceedings (Lecture Notes in Computer Science). Berlin Heidelberg: Springer; 2006 Presented at: Pervasive Computing: 4th International Conference; May 7-10, 2006; Dublin, Ireland p. 366-382.

110. Gu T, Wu Z, Tao X, Pung HK, Lu J. epSICAR: An Emerging Patterns based approach to sequential, interleaved and Concurrent Activity Recognition. 2009 Presented at: 7th Annual IEEE International Conference on Pervasive Computing Communications; March 9-13, 2009; Galveston, TX. [doi: 10.1109/PERCOM.2009.4912776]

111. Horrocks I, Patel-Schneider P, van Harmelen F. From SHIQ and RDF to OWL: the making of a Web Ontology Language. Web Semantics: Science, Services and Agents on the World Wide Web 2003 Dec;1(1):7-26. [doi: 10.1016/j.websem.2003.07.001]

112. Chen L, Nugent C. Ontology‐based activity recognition in intelligent pervasive environments. Int J of Web Info Systems 2009 Nov 20;5(4):410-430. [doi: 10.1108/17440080911006199]

113. Riboni D, Bettini C. COSAR: hybrid reasoning for context-aware activity recognition. Pers Ubiquit Comput 2010 Aug 14;15(3):271-289. [doi: 10.1007/s00779-010-0331-7]

114. Pal-SPOT project. URL: http://everywarelab.di.unimi.it/palspot [accessed 2015-06-09] [WebCite Cache ID 6Z9zRw7nG]

115. Lakiotaki K, Delias P, Sakkalis V, Matsatsinis N. User profiling based on multi-criteria analysis: the role of utility functions. Oper Res Int J 2008 Nov 20;9(1):3-16. [doi: 10.1007/s12351-008-0024-4]

116. Goldberg LR. The structure of phenotypic personality traits. Am Psychol 1993 Jan;48(1):26-34. [Medline: 8427480]

117. McCrae R, Costa PT. Toward a new generation of personality theories: Theoretical contexts for the five-factor model. In: Wiggins JS, editor. The five-factor model of personality: Theoretical perspectives. New York, NY: Guilford; 1996:51-87.

118. Eysenck HJ. Personality and prediction - principles of personality-assessment. In: Wiggins J, editor. Personality and individual differences. USA: Wesley Publishing; 1973.

119. Arya A, Jefferies LN, Enns JT, DiPaola S. Facial actions as visual cues for personality. Comp Anim Virtual Worlds 2006 Jul;17(3-4):371-382. [doi: 10.1002/cav.140]

120. Abowd GD, Dey AK, Brown PJ, Davies N, Smith M, Steggles P. Towards a better understanding of context and context-awareness. 1999 Sep 27 Presented at: 1st International Symposium on Handheld and Ubiquitous Computing; September 27-29, 1999; Karlsruhe, Germany p. 304-307.

121. Lieberman H, Fry C, Weitzman L. Exploring the Web with reconnaissance agents. Commun. ACM 2001;44(8):69-75. [doi: 10.1145/381641.381661]

122. Brusilovsky P, Sosnovsky S, Shcherbinina O. User modeling in a distributed e-learning architecture. 2005 Presented at: 10th International Conference on User Modeling; 2005; Edinburgh, UK p. 387-391.

123. Chen CM, Hsieh YL. Mining learner profile utilizing association rule for common learning misconception diagnosis. 2005 Presented at: Fifth IEEE International Conference on Advanced Learning Technologies; 2005; Kaohsiung, Taiwan p. 588-592.

124. Charniak E. A Bayesian model of plan recognition. Artificial Intelligence 1993 Nov;64(1):53-79. [doi: 10.1016/0004-3702(93)90060-O]

125. Jensen F. Bayesian Networks and Decision Graphs (Information Science and Statistics). Berlin: Springer; 2007.

126. Siadaty MS, Knaus WA. Locating previously unknown patterns in data-mining results: a dual data- and knowledge-mining method. BMC Med Inform Decis Mak 2006;6:13 [FREE Full text] [doi: 10.1186/1472-6947-6-13] [Medline: 16522200]

127. Kolodner JL. Improving Human Decision Making through Case-Based Decision Aiding. AI Magazine 1991:52-68.

128. Geissler J. Informations to patients debate - as if the internet was still a walled garden. 2010 Apr 21. URL: http://www.rarediseaseblogs.net/information-to-patients-debate-2010-as-if-the-internet-was-still-a-walled-garden/ [accessed 2015-06-09] [WebCite Cache ID 6Z9zthOaB]

129. Genitsaridi I, Kondylakis H, Koumakis L, Marias K, Tsiknakis M. Evaluation of personal health record systems through the lenses of EC research projects. Comput Biol Med 2015 Apr;59:175-185. [doi: 10.1016/j.compbiomed.2013.11.004] [Medline: 24315661]

130. Sunyaev A, Chornyi D, Mauro C, Krcmar H. Evaluation Framework for Personal Health Records: Microsoft HealthVault Vs. Google Health. 2010 Presented at: 43rd Hawaii International Conference on System Sciences (HICSS); 5-8 Jan. 2010; Honolulu, HI p. 1-10. [doi: 10.1109/HICSS.2010.192]

131. Baudendistel I, Winkler E, Kamradt M, Längst G, Eckrich F, Heinze O, et al. Personal electronic health records: understanding user requirements and needs in chronic cancer care. J Med Internet Res 2015;17(5):e121 [FREE Full text] [doi: 10.2196/jmir.3884] [Medline: 25998006]

132. Ratzan S, Parker R. National Library of Medicine Current Bibliographies in Medicine: Health Literacy. USA: National Library of Medicine; 2000.

133. Fogg B. Persuasive technology. Ubiquity 2002 Dec 01;2002(December):2. [doi: 10.1145/764008.763957]

134. Kaptein M, De Ruyter B, Markopoulos P, Aarts E. Adaptive Persuasive Systems. ACM Trans Interact Intell Syst 2012 Jun 01;2(2):1-25. [doi: 10.1145/2209310.2209313]

135. Spanakis EG, Santana S, Ben-David B, Marias K, Tziraki C. Persuasive technology for healthy aging and wellbeing. 2014 Presented at: 4th International Conference on Wireless Mobile Communication and Healthcare; Nov. 3-5, 2014; Athens, Greece.

XSL•FO

RenderX

136. With-Me-Project, The European Platform to Promote Healthy Lifestyle and improve care through a Personal Persuasive Assistant. URL: http://www.with-me-project.eu/ [accessed 2015-06-09] [WebCite Cache ID 6Z9zaryOo]

137. van Gemert-Pijnen J, Nijland N, van LM, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]

138. Sakkalis V. Applied strategies towards EEG/MEG biomarker identification in clinical and cognitive research. Biomark Med 2011 Feb;5(1):93-105. [doi: 10.2217/bmm.10.121] [Medline: 21319971]

139. Emotiv brain EEG. URL: http://www.emotiv.com [accessed 2015-06-09] [WebCite Cache ID 6Z9zxvzAE]

140. Neyrosky ECG. EEG biosensor URL: http://neurosky.com [accessed 2015-06-09] [WebCite Cache ID 6ZA08iFXl]

141. Sakkalis V, Zervakis M, Micheloyannis S. Significant EEG features involved in mathematical reasoning: evidence from wavelet analysis. Brain Topogr 2006;19(1-2):53-60. [doi: 10.1007/s10548-006-0012-z] [Medline: 16988882]

142. Chiarugi F, Iatraki G, Christinaki E, Manousos D, Giannakakis G, Pediaditis M, et al. Facial signs and psycho-physical status estimation for well-being assessment. 2014 Presented at: 7th IEEE International Conference on Health Informatics; 2014; Angers, France.

143. Sakkalis V. Review of advanced techniques for the estimation of brain connectivity measured with EEG/MEG. Comput Biol Med 2011 Dec;41(12):1110-1117. [doi: 10.1016/j.compbiomed.2011.06.020] [Medline: 21794851]

144. Ktonas PY, Golemati S, Xanthopoulos P, Sakkalis V, Ortigueira MD, Tsekou H, et al. Time-frequency analysis methods to quantify the time-varying microstructure of sleep EEG spindles: possibility for dementia biomarkers? J Neurosci Methods 2009 Dec 15;185(1):133-142. [doi: 10.1016/j.jneumeth.2009.09.001] [Medline: 19747507]

145. Dekker MKJ, Sitskoorn MM, Denissen AJM, van Boxtel GJM. The time-course of alpha neurofeedback training effects in healthy participants. Biol Psychol 2014 Jan;95:70-73. [doi: 10.1016/j.biopsycho.2013.11.014] [Medline: 24321361]

146. Sakkalis V, Oikonomou T, Pachou E, Tollis I, Micheloyannis S, Zervakis M. Time-significant wavelet coherence for the evaluation of schizophrenic brain activity using a graph theory approach. Conf Proc IEEE Eng Med Biol Soc 2006;1:4265-4268. [doi: 10.1109/IEMBS.2006.260680] [Medline: 17946233]

147. Murias M, Webb SJ, Greenson J, Dawson G. Resting state cortical connectivity reflected in EEG coherence in individuals with autism. Biol Psychiatry 2007 Aug 1;62(3):270-273 [FREE Full text] [doi: 10.1016/j.biopsych.2006.11.012] [Medline: 17336944]

148. Sakkalis V, Tsiaras V, Zervakis M, Tollis I. Optimal brain network synchrony visualization: application in an alcoholism paradigm. Conf Proc IEEE Eng Med Biol Soc 2007;2007:4285-4288. [doi: 10.1109/IEMBS.2007.4353283] [Medline: 18002949]

149. Kondylakis H, Koumakis L, Genitsaridi E, Tsiknakis M, Marias K, Pravettoni G, et al. IEmS: A collaborative environment for patient empowerment. 2012 Presented at: IEEE International Conference on BioInformatics and BioEngineering (BIBE); Nov. 11-13, 2012; Larnaca.

150. Karatzanis I, Kontogiannis V, Spanakis EG, Chiarugi F, Fursse J, Jones RW. Empowering patients through a patient portal for an improved diabetes management. Wireless Mobile Communication and Healthcare, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2013;61:398-405. [doi: 10.1007/978-3-642-37893-5_44]

151. Spanakis EG, Chiarugi F, Kouroubali A, Spat S, Beck P, Asanin S, et al. Diabetes management using modern information and communication technologies and new care models. Interact J Med Res 2012;1(2):e8 [FREE Full text] [doi: 10.2196/ijmr.2193] [Medline: 23612026]

152. Kartakis S, Sakkalis V, Tourlakis P, Zacharioudakis G, Stephanidis C. Enhancing health care delivery through ambient intelligence applications. Sensors (Basel) 2012;12(9):11435-11450 [FREE Full text] [doi: 10.3390/s120911435] [Medline: 23112664]

153. Piette JD, Striplin D, Marinec N, Chen J, Trivedi RB, Aron DC, et al. A Mobile Health Intervention Supporting Heart Failure Patients and Their Informal Caregivers: A Randomized Comparative Effectiveness Trial. J Med Internet Res 2015;17(6):e142 [FREE Full text] [doi: 10.2196/jmir.4550] [Medline: 26063161]

154. Spanakis EG, Lelis P, Chiarugi F, Chronaki C. R&D challenges in developing an ambient intelligence eHealth platform. 2005 Presented at: 3rd European Medical and Biological Engineering Conference (EMBEC 2005); November 20-25, 2005; Prague, Czech Republic p. 1727-1983.

155. Barua M, Lu R, Shen X. Health-Post: A Delay-Tolerant Secure Long-Term Health Care Scheme in a Rural Area. In: GLOBECOM 2011. 2011 Presented at: Global Telecommunications Conference; Dec. 5-9, 2011; Houston, TX.

156. Spanakis EG, Voyiatzis AG. DAPHNE: A Disruption-Tolerant Application Proxy for e-Health Network Environment. In: MobiHealth 2012. Berlin, Germany: Springer; 2012 Nov 21 Presented at: 3rd International Conference on Wireless Mobile Communication and Healthcare; November 21-23, 2012; Paris, France p. 88-95. [doi: 10.1007/978-3-642-37893-5_10]

157. Genitsaridi I, Kondylakis H, Koumakis L, Marias K, Tsiknakis M. Towards Intelligent Personal Health Record Systems: Review, Criteria and Extensions. Procedia Computer Science 2013;21:327-334. [doi: 10.1016/j.procs.2013.09.043]

158. Kotz D. A threat taxonomy for mHealth privacy. In: COMSNETS 2011. 2011 Presented at: IEEE Third International Conference on Communication Systems and Networks; Jan. 4-8, 2011; Bangalore.

XSL•FO

RenderX

159. Mare S, Sorber J, Shin M, Cornelius C, Kotz D. Adapt-lite: Privacy-aware, secure, and efficient mhealth sensing. 2011 Presented at: 10th Annual ACM Workshop on Privacy in the Electronic Society; Oct. 17-21, 2011; Chicago, IL.

160. Spanakis EG, Tsiknakis MN. Delivery of secure eHealth services and communications. In: ICTSM2012. 2012 Presented at: International Conference on Telecommunication Systems, Modeling and Analysis; May 24-26, 2012; Prague, Czech Republic.

161. Deyan C, Hong Z. Data security and privacy protection issues in cloud computing. 2012 Presented at: International Conference on Computer Science Electronics Engineering (ICCSEE); March 25, 2012; Hangzhou p. 647-651. [doi: 10.1109/ICCSEE.2012.193]

162. Saarinen J. IT news. 2012 Aug 07. UK health trust fined for privacy breach URL: http://www.itnews.com.au/News/311079,ukhealth-trust-fined-for-privacy-breach.aspx [accessed 2015-06-09] [WebCite Cache ID 6Z9ynrOha]

163. SeongHan S, Kobara K, Imai H. A secure public cloud storage system. 2011 Presented at: International Conference on Internet Technology and Secured Transactions (ICITST); 2011; Abu Dhabi, United Arab Emirates.

164. Minqi Z, Rong Z, Wei X, Weining Q, Aoying Z. Security and privacy in cloud computing: a survey. 2010 Presented at: Sixth International Conference on Semantics KnowledgeGrid (SKG); 2010; Beijing p. 105-112.

165. van't Noordende GJ, Olabarriaga SD, Koot MR, de Laat CTA. A Trusted Data Storage Infrastructure for Grid-Based Medical Applications. In: CCGRID '08. 2008 Presented at: 8th IEEE International Symposium on Cluster Computing and the Grid; May 19-22, 2008; Lyon, France.

166. Kuner C. European data protection law: corporate compliance and regulation. Oxford: Oxford University Press; 2007.

167. Department of Health and Human Services. HIPAA. URL: http://www.hhs.gov/ocr/privacy/index.html [accessed 2015-06-09] [WebCite Cache ID 6Z9ygLj02]

168. IEEE Recommended Practice for Architectural Description of Software-Intensive Systems. URL: http://cabibbo.dia.uniroma3.it/ids/altrui/ieee1471.pdf [accessed 2016-05-12] [WebCite Cache ID 6hRxhNgPA]

169. Kahn EB, Ramsey LT, Brownson RC, Heath GW, Howze EH, Powell KE, et al. The effectiveness of interventions to increase physical activity. A systematic review. Am J Prev Med 2002 May;22(4 Suppl):73-107. [Medline: 11985936]

170. Sakkalis V, Sfakianakis S, Marias K. Bridging social media technologies and scientific research: an exemplary platform for VPH modelling. In: MobiHealth 2012. 2012 Presented at: 3rd International ICST Conference on Wireless Mobile Communication and Healthcare; November, 2012; Paris, France.

171. Portnoy DB, Scott-Sheldon LAJ, Johnson BT, Carey MP. Computer-delivered interventions for health promotion and behavioral risk reduction: a meta-analysis of 75 randomized controlled trials, 1988-2007. Prev Med 2008 Jul;47(1):3-16 [FREE Full text] [doi: 10.1016/j.ypmed.2008.02.014] [Medline: 18403003]

Abbreviations

ACSM: American College of Sports Medicine
ASTM: American Society for Testing and Materials
CEN-TC: Comité Européen de Normalisation - Technical Committee
EHR: electronic health record
GPS: global positioning system
HDP: health device profile
HL7: Health Level Seven
ICT: information and communication technology
IT: information technology
NCD: non-communicable chronic diseases
SOA: service-oriented architecture
SODA: service-oriented device architecture

[XSL•FO]

RenderX

<u>Original Paper</u>

Changing Mental Health and Positive Psychological Well-Being Using Ecological Momentary Interventions: A Systematic Review and Meta-analysis

Anke Versluis[1], MSc; Bart Verkuil[2], PhD; Philip Spinhoven[2,3], PhD; Melanie M van der Ploeg[1], MSc; Jos F Brosschot[1], PhD

[1]Health, Medical and Neuropsychology Unit, Institute of Psychology, Leiden University, Leiden, Netherlands
[2]Clinical Psychology Unit, Institute of Psychology, Leiden University, Leiden, Netherlands
[3]Department of Psychiatry, Leiden University Medical Center, Leiden, Netherlands

Corresponding Author:
Anke Versluis, MSc
Health, Medical and Neuropsychology Unit
Institute of Psychology
Leiden University
Wassenaarseweg 52
Leiden, 2333 AK
Netherlands
Phone: 31 715276343
Fax: 31 71 527 3619
Email: a.versluis@fsw.leidenuniv.nl

Abstract

Background: Mental health problems are highly prevalent, and there is need for the self-management of (mental) health. Ecological momentary interventions (EMIs) can be used to deliver interventions in the daily life of individuals using mobile devices.

Objectives: The aim of this study was to systematically assess and meta-analyze the effect of EMI on 3 highly prevalent mental health outcomes (anxiety, depression, and perceived stress) and positive psychological outcomes (eg, acceptance).

Methods: PsycINFO and Web of Science were searched for relevant publications, and the last search was done in September 2015. Three concepts were used to find publications: (1) mental health, (2) mobile phones, and (3) interventions. A total of 33 studies (using either a within- or between-subject design) including 43 samples that received an EMI were identified (n=1301), and relevant study characteristics were coded using a standardized form. Quality assessment was done with the Cochrane Collaboration tool.

Results: Most of the EMIs focused on a clinical sample, used an active intervention (that offered exercises), and in over half of the studies, additional support by a mental health professional (MHP) was given. The EMI lasted on average 7.48 weeks (SD=6.46), with 2.80 training episodes per day (SD=2.12) and 108.25 total training episodes (SD=123.00). Overall, 27 studies were included in the meta-analysis, and after removing 6 outliers, a medium effect was found on mental health in the within-subject analyses (n=1008), with g=0.57 and 95% CI (0.45-0.70). This effect did not differ as function of outcome type (ie, anxiety, depression, perceived stress, acceptance, relaxation, and quality of life). The only moderator for which the effect varied significantly was additional support by an MHP (MHP-supported EMI, g=0.73, 95% CI: 0.57-0.88; stand-alone EMI, g=0.45, 95% CI: 0.22-0.69; stand-alone EMI with access to care as usual, g=0.38, 95% CI: 0.11-0.64). In the between-subject studies, 13 studies were included, and a small to medium effect was found (g=0.40, 95% CI: 0.22-0.57). Yet, these between-subject analyses were at risk for publication bias and were not suited for moderator analyses. Furthermore, the overall quality of the studies was relatively low.

Conclusions: Results showed that there was a small to medium effect of EMIs on mental health and positive psychological well-being and that the effect was not different between outcome types. Moreover, the effect was larger with additional support by an MHP. Future randomized controlled trials are needed to further strengthen the results and to determine potential moderator variables. Overall, EMIs offer great potential for providing easy and cost-effective interventions to improve mental health and increase positive psychological well-being.

XSL•FO
RenderX

Introduction

One in every 3 individuals worldwide will be affected by one or more mental health problems during their lives [1]. Yet, only a small portion of those individuals is receiving help for their problems (with numbers varying from 7% to 25% in industrialized countries) [2,3]. To help those in need, new strategies for enhancing access to and quality of care are needed, and this is recognized in a new policy of the World Health Organization [4]. This newly introduced policy requests methods to increase self-management or self-care of health by, for instance, using electronic and mobile devices. In line with this, Wanless [5] argues that health care productivity can be increased using self-care and that this can have cost-effective benefits. All in all, there appears to be a future for the self-management of (mental) health.

One method that can be used to enhance health self-management is ecological momentary interventions (EMIs) [6]. The key to these interventions is that they can be tailored to the individual and be implemented in real time (ie, daily life). Mobile or electronic devices can be used to provide these interventions in the daily lives of individuals. With a Web-based survey, Proudfoot et al [7] showed that 76% of the general population is interested in using mobile technology for either self-monitoring or self-management of health (ie, if the service was free). Using EMIs has numerous advantages such as the ability to reach large populations at lower costs [8,9].

Training people in situ could be highly relevant for learning new, healthy behaviors, considering that people under stress typically switch from *goal-directed behavior* to *habit behavior* [10-13]. In other words, when a person experiences stress, that person is more likely to rely on the "old" behavior routine than display the newly learned behavior routine. In line with this, it might make more sense to learn a new behavioral routine in daily life compared with an artificial surrounding (eg, the therapist's office) that generally does not resemble daily life. Indeed, research shows that although new behaviors can be effectively learned in artificial surroundings, this knowledge does not always generalize to real-life settings [14]. According to Neal et al [15], this is understandable, given that the association between context and the maladaptive behavior may still be in place after traditional treatment. As a consequence, the context (eg, setting or time of day) can still trigger the maladaptive behavior. Therefore, EMIs may provide a more effective way to train people in daily life than conventional treatment, by training people in the very context in which the maladaptive behavior occurs. As a result, this could lead to the (faster) formation of a new and more adaptive association between context and behavior.

Given that the number of worldwide mobile phone users is immense and continues to expand [16], it is not surprising that EMI is considered to be the future for therapeutic interventions

[17]. Numerous authors highlight that EMI is a relatively new research field, and that the field is constantly evolving due to improvements in mobile technology [17-19]. It is therefore important to know the current state of affairs in this field. Current reviews suggest that EMIs can be effective, but these reviews are limited for different reasons. First, some reviews focus on a specific intervention [20] or on a specific target population [21]. Second, their sole or main focus is the effect of EMIs on health behaviors (eg, physical activity, smoking cessation, diabetes management) and not mental health [18,22,23]. Third, the current reviews are outdated, especially considering the developmental pace of EMIs (eg, [19]). A more recent review has been conducted by Donker et al [24]; however, it included only studies that investigated directly downloadable apps. This substantially limited the number of included studies (n=8). Fourth, the effect of EMIs on positive psychological well-being (eg, relaxation, acceptance) has not yet been reviewed, although these outcome types have been included as dependent variables in previous studies [25,26]. Considering that a person's well-being is not equal to the absence of disease and is associated with increased positive cognitions and even physical health, it is important to also study these positive experiences [27]. To conclude, an up-to-date comprehensive overview or a meta-analysis of the effect of EMIs on mental health, including positive health outcomes, is missing.

This systematic review and meta-analysis therefore attempts to expand the current knowledge by including both mental health outcomes (ie, perceived stress, anxiety, or depressive symptoms) and positive psychological outcomes (eg, positive affect or acceptance). For this quantitative analysis, randomization and the presence of a control group were optional. Although the absence of randomization and the lack of a control group may weaken the design and thus the ensuing conclusions, these criteria are necessary to ensure that the presented overview of EMI studies is complete. This is considered critical because an extensive overview is currently lacking. It should be noted that study design was used in the moderator analyses.

Considering that the access to care needs improvement and EMIs can be used for this, it is important to investigate for whom these technologies can be appropriate and what EMI characteristics are associated with increased effects. Therefore, potentially promising moderators of effect size were investigated. Specifically, sample, type of training, how the training was triggered (ie, automatically or on-demand), support of mental health professional (MHP), and dosage were included because these can be considered key intervention components [28]. Including moderators allows us, for example, to investigate whether an EMI in its own right is effective or whether additional support by an MHP is necessary to accomplish change. In addition, the design of the study, sample size, and the quality of the study were studied to determine whether the effect size varied as a function of study characteristics. In short, we examined whether mobile technology provides an effective

XSL•FO

RenderX

platform for mental health interventions and under which circumstances.

Methods

The preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines were followed [29].

Search Strategies

To find relevant publications concerning EMIs that target mental health, a database search was conducted in both PsycINFO and Web of Science (Core Collection). The search strings that were used consisted of 3 groups of words, namely words related to: (1) mental health, (2) mobile phones, and (3) interventions. See Multimedia Appendix 1 for the complete search strings. In both the databases, the search was limited to English publications that were peer reviewed. The search strategy was not restricted based on publication year as we aimed to provide a comprehensive overview of how mobile technology can be used to improve mental health. Naturally, the technologies that are used in more recent publications may be more advanced compared with earlier publications, but the idea of repeatedly training people in their daily lives is equal in older and newer publications. The last search was conducted on September 17, 2015. In addition, 2 other search strategies were used. First, the reference lists of previous reviews in the field of EMI were screened for relevant publications. Second, the reference lists of our primary selected papers were examined.

To ensure that no relevant publications were missed with the aforementioned search strategies, an extra search with a similar search string was conducted in the PubMed database on November 2, 2015. This resulted in 3505 publications, and the first 10% was screened to determine whether potentially relevant studies had been missed. However, no relevant publications—that had not already been identified in the other databases—were found, indicating that the used search strategies were sufficient.

Study Selection

Titles and abstracts of publications were first screened for eligibility, and if insufficient information was described in the abstract, the full-text papers were obtained. When a full-text paper was not available, a request was sent to the authors. A number of inclusion criteria were used for both within- and between-subject studies, which were established by authors AV, BV, and JB. First, publications were included when an EMI was studied (eg, via smartphone or personal digital assistant)—either as a stand-alone intervention or in combination with other treatment components. Second, the EMI should be automated and operated independently from a therapist. Thus, studies were excluded when the therapist administered the therapy—for instance—via mobile phone or conference call. This criterion was chosen because of our interest in how new technologies could be used to deliver *cost-effective* treatments in daily life, which precluded those requiring comparatively conventional therapist's efforts. Third, a mental health–related outcome should be targeted (eg, anxiety, depression, or positive psychological well-being and not a health-related outcome such as physical activity). Fourth, the EMI should be studied in an

ambulatory setting and not in standard therapy sessions. Publications were excluded if a mental health–related outcome was included, but the training was not directly focused on improving mental health (eg, psychoeducation for health behaviors or hypertension management). Moreover, studies that did not discuss post-intervention outcome data, without a baseline measure, methodological papers, case studies, reviews, non–peer-reviewed papers, and non-English papers were excluded. Three publications were additionally excluded because the samples were already discussed in other, already included publications. If a study included a control group—in addition to the group that received the EMI—it was coded as a between-subject study (see *Coding* for further details). The screening was conducted by author AV, and uncertainty about the potential inclusion or exclusion of a paper was resolved with authors BV and JB.

Coding

To collect the relevant study characteristics from each publication, a standardized form was used. Using this form, the following data were collected: (1) first author and publication year, (2) design, (3) sample characteristics (clinical characteristics, age, gender, and sample size), (4) outcome type, (5) information on the EMI (training type, training trigger, number of training episodes, and whether training was supported by an MHP), and (6) type of control condition and sample size. When a publication reported on more than 1 EMI, information was extracted separately for each described EMI, and all EMIs were included separately in the within-subject analyses. For the between-subject analyses, however, only 1 EMI was included thereby ensuring that each participant is represented only once in the analyses [30]. The EMI that was included in the between-subject analyses was the most "complete" intervention. In the case of Grassi et al [25], the Vnar intervention was chosen because it included both video and audio components compared with a video- or audio-only intervention. For both the studies by Repetto et al [31] and Pallavicini et al [32], the virtual reality intervention with biofeedback was chosen above the intervention using only virtual reality.

In the meta-analysis, the primary outcome of interest was "mental health." Mental health encompasses an anxiety, depression, or stress outcome. Per publication, a set of guidelines was used to determine which specific questionnaire was used to represent this primary outcome. If a study reported 1 primary outcome, this measure was chosen as an indicator of mental health. When no or multiple primary outcomes were defined, a measure was chosen that was most likely to be affected given the aim of the training. For example, if the training focused on reducing anxiety, then, an anxiety questionnaire was preferred over a questionnaire measuring depression. In this process of selecting questionnaires, comprehensive questionnaires were chosen over restricted questionnaires (if there was such a choice), and the most valid questionnaire was chosen (idem). In addition to the coding of the primary outcome for each publication, the different outcome types per study were also coded. Thus, all questionnaires measuring anxiety, depression, perceived stress, and positive psychological well-being outcomes were listed per publication. A questionnaire was considered to represent positive psychological well-being, when it specifically

identified positive emotions or processes that were targeted with the intervention. The only positive psychological well-being outcomes that were identified in the publications were acceptance, feelings of relaxation and quality of life; positive affect, for instance, was not studied in the included publications. By listing all the questionnaires that measured mental health and positive psychological well-being, it was possible to examine whether the effectiveness of EMI differed per outcome type (eg, anxiety or depression).

With regard to the information on the EMI, it was reported whether the training was active or passive. A training was labeled as active when participants had to carry out an exercise, for instance, a relaxation exercise [33]. In contrast, a passive training supplied information to the participants (eg, suggestions or tips) but did not require an immediate action from the participant. For example, participants are given messages that would support self-management [34]. Furthermore, when a trigger (using the EMI device) reminds participants to do the training at a specific moment, the training was coded as "triggered." If participants could do the training whenever they preferred, the triggering of the training was said to be "on-demand." Moreover, it was reported whether the EMI was used as a stand-alone intervention (coded as stand-alone EMI) or was part of a treatment package and was thus supported by an MHP (coded as MHP-supported EMI). This treatment package could consist of either an EMI in combination with therapy (eg, group therapy or exposure therapy) or an EMI with continued feedback (eg, feedback on homework exercises or messages to improve adherence). An introductory or kickoff session at the start of the intervention was not coded as support. When the effect of an EMI was studied in a population that had access to care as usual (eg, inpatient or outpatient setting), but this (additional) care was not the focus of the study or was not specifically related to the EMI, the EMI was coded as a stand-alone intervention in combination with care as usual. However, these studies often did not specify whether this available care was used by individuals or what this care specifically entailed. Finally, if a study included a control condition and was therefore eligible for the between-subject analyses, the type of control condition was reported (waitlist, placebo, or active treatment). Specifically, if more than 1 control condition was used, a placebo condition was chosen over a waitlist condition, and an active treatment control condition was chosen over both the placebo and waitlist condition. When multiple active treatment control conditions were included in the study, the condition was chosen that had the closest resemblance with the EMI condition, but without its "target ingredient." This way it was possible to more precisely determine the added value of mobile technology when delivering interventions. Although it is possible to include all reported control conditions using multiple pairwise comparisons (eg, intervention group vs placebo and intervention group vs waitlist), this yields problems in the analyses as the same group is overrepresented (eg, twice). Therefore, in the case of the studies of Kenardy et al [35] and Newman et al [36], the 6-session cognitive behavioral therapy (CBT) was chosen to represent the control condition because it better resembled the EMI condition (6 sessions of computer-assisted CBT) compared with the 12-session CBT condition. Review author (AV)

extracted all the relevant study characteristics from the included publications. To check the inter-rater reliability, a second reviewer (MvdP) assessed data from a subset of the selected papers (ie, 20%) [37]. For the nominal variables, the average Cohen's kappa was .86 indicating strong agreement between the 2 raters. The other variables had an 88% (37/42) agreement, which demonstrates a high consistency among raters.

Quality Assessment

The risk of bias in individual studies was assessed using the Cochrane Collaboration tool [38]. This assessment tool uses 6 different domains for determining the quality of randomized trials: (1) selection bias concerns the method used to generate and conceal the allocation sequence (random sequence generation and allocation concealment, respectively); (2) performance bias deals with ways in which participants and personnel are blinded from knowing condition allocation; (3) detection bias relates to measures that are taken to blind the outcome assessment from knowledge of which intervention participants received; (4) attrition bias refers to whether the study attrition and exclusions from analysis are reported; (5) reporting bias is whether selective outcome reporting is examined and discussed; (6) other bias refers to any other problems or concerns that are not addressed by previous points. For each publication, the domains are rated with either a "high" or "low" risk. If insufficient information is provided in the paper, then, the level of risk is labeled "unclear." Higgins et al [38] argues that within the domain "other bias," the sources of bias should be prespecified. In this case, no other biases were specified in advance; therefore, this domain was omitted from the current quality assessment.

The quality assessment was done by the first author (AV), and a 20% sample was assessed by a second reviewer (MvdP). Inter-rater reliability, as assessed with Cohen's kappa, indicated that there was moderate agreement between raters (ie, average kappa of .69).

Data Analysis

Hedges' *g* was used as an estimate of the effect size. This estimate was calculated using the mean, SD, and sample size at post-intervention as reported in the paper or as based on contact with the authors. Moreover, to compute an effect, a correlation coefficient is needed that represents the correlation between the repeated measures of the outcome parameter. As this within-subject correlation was rarely reported, the correlation was set at .50 for all studies [39]. For interpreting the effect size, the guidelines for Cohen's *d* were used because they are approximately compatible [40]. According to these guidelines, a value of 0.20 is small, 0.50 is medium, and 0.80 is large. Effect sizes are based on a random effect model because we expect the real effect to differ between studies.

To estimate the effect of EMI from pre intervention to postintervention, analyses were first run with all within-subject data. Furthermore, to determine whether this effect differed from a control condition, between-subject analyses were run. In both the within- and between-subject analyses, it was determined whether there was an effect on the primary outcome "mental health" (as measured with a single questionnaire).

Second, it was investigated whether the effect differed per outcome type. That is, was the effect of EMI different for anxiety, depression, perceived stress, or positive psychological outcomes (acceptance, relaxation, and quality of life). To determine the effectiveness per outcome type, all relevant outcome types per publication were included in the analysis. When a study used multiple questionnaires to assess an outcome type (eg, anxiety), an overall mean was created by combining these different questionnaires. By combining multiple questionnaires per study, the data are unlikely to be independent, and this increases the type II error. Therefore, these analyses are only used to explore whether there are potential differences in effects between the outcome types. In addition, for the primary outcome "mental health," subgroup analyses are done to determine whether the effect differed as a function of design (randomized controlled trial [RCT] or pre-post), sample (healthy or clinical), age, gender, sample size, training type (active or passive), training trigger (triggered, on-demand, or unspecified), daily training episodes (number), total training episodes (number), support by MHP (stand-alone EMI, MHP-supported EMI, or stand-alone EMI with access to care as usual), and quality assessment (0-6). Year of publication was not included as a moderator because there was little variation in this variable (ie, 25 of the 32 publications were published in 2010 or later). Moreover, type of control condition was not included as a moderator because only 13 studies had a between-subject design.

As a measure of heterogeneity, the Q and I^2 statistics were used. A significant Q-statistic indicates that there is variation in the true effect size, and I^2 reflects the amount of real variance—specifically, values of 25%, 50%, and 75% can be considered small, medium, and large values, respectively [41]. Moreover, the risk for publication bias was examined using different techniques [30]. First, the distribution in the funnel plot was visually inspected as a preliminary indication for publication bias. This plot represents the effect size against the standard error of the study. Generally, studies with a large sample size are represented at the top of the plot around the mean, and studies with a smaller sample size are located at the bottom of the plot with a wider distribution around the mean. In the case of publication bias, studies with a small sample size are more likely to fall to the right of the mean (indicating a positive effect size). In other words, when the distribution of studies becomes asymmetrical, there is indication for publication

bias. To quantify the amount of bias, the Egger's test of intercept was used. In this approach, the amount of bias is captured in the intercept value, and a significant intercept indicates that there is significant publication bias. Furthermore, to correct for the missing studies (to the left of the mean), a Duval and Tweedie's trim and fill method was used. This method calculates where missing studies were most likely to fall and adds these studies to the analysis. The recomputed effect size and CI are thereby corrected for the missing studies and is assumed to be unbiased [30].

Outliers were identified using the value of the standardized residual in both the within- and between-subject analyses. Studies whose standardized residual was significant (values ± 1.96) were excluded from the analyses.

The software Comprehensive Meta-Analysis version 3.3.070 (Biostat) was used for all the described analyses including the calculation of effect sizes with 95% CIs. The forest plots were made using the metaphor package in R (version 3.0.3) [42].

Results

A total of 2611 publications were identified with the search strategies after removing duplicates (see Figure 1) [29]. After screening the titles and abstracts, 127 full-text publications were screened for eligibility. Most of these publications were excluded because no (mobile phone) intervention was studied, the intervention was not automated (ie, not independent from therapist), or no outcome data were discussed (methodological paper). A total of 32 publications were considered relevant and were included in the analysis (see Tables 1 and 2). In these 32 publications, 33 different studies were reported using 43 samples that received an EMI (n=1301). The included study by Huffziger et al [26] was technically an ecological momentary assessment study (with an experimental manipulation) and not an EMI. However, considering that the manipulation that was used (mindfulness attention induction) can be seen as an intervention, the study was included.

For the meta-analysis, 5 publications were excluded because no means and SDs to calculate the effect size were reported or obtained after contacting the authors [43-47]. Therefore, 27 publications (27 studies) with 33 samples that received an EMI were included in the meta-analysis (n=1156).

XSL·FO
RenderX

Table 1. Characteristics of the ecological momentary intervention studies (part 1).

Study[a]	Design[b]	Sample	Age (years)	Gender (% female)	n[c]	Mental Health Measure[d]	Outcome type(s)
Included in meta-analysis							
Agyapong et al, 2012[e]	RCT	Clinical	48.00	54	24	BDI	Depression
Ahtinen et al, 2013	Prepost	Healthy	—	60	14	Stress single-item	Stress Acceptance Quality of life
Aikens et al. 2015[f] (all pooled subjects)	Prepost	Clinical	51.40	79	221	PHQ-8	Depression
Askins et al, 2009	RCT	Healthy	36.30	100	64	POMS	Depression
Ben-Zeev et al, 2014	Prepost	Clinical	45.90	39	32	BDI	Depression
Burns et al, 2011[e]	Prepost	Clinical	37.40	88	7	GIDS-c	Depression Anxiety
Carissoli et al, 2015	RCT	Healthy	38.11	57	20	MSP	Stress
Dagöö et al. 2014[g] (mCBT)	RCT	Clinical	34.70	48	24	LSAS-SR	Depression Anxiety Quality of life
Dagöö et al, 2014[g] (mIPT)	RCT	Clinical	39.08	56	19	LSAS-SR	Depression Anxiety Quality of life
Depp et al, 2015	RCT	Clinical	46.90	54	41	MADRS	Depression
Enock et al. 2014	RCT	Clinical	34.80	48	120	SIAS	Depression Anxiety
Granholm et al, 2012	Prepost	Clinical	48.70	31	41	BDI	Depression
Grassi et al, 2007 (Vnar)	Prepost[h]	Healthy	23.27	50	30	STAI-state	Anxiety Relaxation
Grassi et al, 2007 (Nnar)	Prepost[h]	Healthy	23.27	50	30	STAI-state	Anxiety Relaxation
Grassi et al, 2007[e] (MP3)	Prepost[h]	Healthy	23.27	50	30	STAI-state	Anxiety Relaxation
Harrison et al, 2011	Prepost	Clinical	38.20	71	28	DASS total score	Depression Anxiety
Huffziger et al, 2013[i]	Prepost	Healthy	22.90	60	46	Valence 2-items	Depression Relaxation
Kenardy et al, 2003[e]	RCT	Clinical	36.80	76	41	Anxiety composite score	Anxiety
Lappalainen et al, 2013	RCT	Clinical	47.10	0	11	GSI	Depression Acceptance Quality of life
Ly et al, 2014[e] (behavioral activation)	RCT	Clinical	36.60	70	36	BDI	Depression Anxiety Acceptance Quality of life
Ly et al, 2014 (mindfulness)	RCT	Clinical	35.60	71	36	BDI	Depression Anxiety Acceptance Quality of life

XSL•FO
RenderX

Study[a]	Design[b]	Sample	Age (years)	Gender (% female)	n[c]	Mental Health Measure[d]	Outcome type(s)
Ly et al, 2012	Prepost	Healthy	29.50	36	11	DASS stress	Depression Anxiety Stress Quality of life
Newman et al, 2014	RCT	Clinical	42.45	55	11	STAI—trait	Anxiety
Newman et al, 1997	RCT	Clinical	38.00	83	9	FQ—total score	Anxiety
Pallavicini et al, 2009 (VRMB)	Prepost[h]	Clinical	41.25	—	4	GAD7	Anxiety
Pallavicini et al, 2009 (VRM)	Prepost[h]	Clinical	48.50	—	4	GAD7	Anxiety
Proudfoot et al, 2013	RCT	Clinical	39.00	70	126	DASS total score	Depression Anxiety Stress
Repetto et al, 2013 (VRMB)	Prepost[h]	Clinical	—	64	7	BAI	Anxiety
Repetto et al, 2013 (VRM)	Prepost[h]	Clinical	—	64	9	BAI	Anxiety
Rizvi et al, 2011	Prepost	Clinical	33.86	82	22	BSI	Depression
Shapiro et al, 2010	Prepost	Clinical	26.30	100	14	BDI	Depression
Watts et al, 2013[e]	RCT	Clinical	41.00	80	10	BDI	Depression Stress
Wenze et al, 2014	Prepost	Clinical	40.86	71	14	QIDS-c	Depression
Not included in meta-analysis							
Gorini et al, 2010 (VRMB)	Prepost[h]	Clinical	—	—	8	BAI	Anxiety
Gorini et al, 2010 (VRM)	Prepost[h]	Clinical	—	—	4	BAI	Anxiety
Grassi et al, 2011 (Vnar)	Prepost[h]	Healthy	20.86	100	15	STAI-state	Anxiety Relaxation
Grassi et al, 2011 (MP3)	Prepost[h]	Healthy	20.86	100	15	STAI-state	Anxiety Relaxation
Preziosa et al, 2009 (Vnar; study 1)	Prepost	Healthy	23.48	100	6	STAI-state	Anxiety Depression
Preziosa et al, 2009 (MP3; study 1)	Prepost	Healthy	23.48	100	6	STAI-state	Anxiety Depression
Preziosa et al, 2009 (study 2)	RCT	Healthy	23.48	50	30	STAI-state	Anxiety Depression Relaxation
Riva et al, 2006	RCT	Healthy	23.82	48	11	STAI-state	Anxiety Depression Relaxation
Zautra et al, 2012 (mindfulness)	RCT	Clinical	54.05	82	25	Depression 3-items	Depression Stress

XSL•FO
RenderX

Study[a]	Design[b]	Sample	Age (years)	Gender (% female)	n[c]	Mental Health Measure[d]	Outcome type(s)
Zautra et al, 2012 (mastery-control)	RCT	Clinical	54.05	82	25	Depression 3-items	Depression stress

[a]Studies are ordered by inclusion in the meta-analysis. Behind the study's year of publication, between brackets, the sample (or condition) that received the ecological momentary intervention was specified; With mCBT: mobile cognitive behavioral therapy; mIPT: mobile interpersonal psychotherapy; MP3: audio only condition; Nnar: video only condition VRMB: virtual reality and mobile condition with biofeedback; VRM: virtual reality with mobile condition; Vnar: video narrative condition.

[b]Design of study is labeled either randomized controlled trial (RCT) or prepost design.

[c]Sample size at post-intervention in the condition receiving the ecological momentary intervention.

[d] The specific questionnaire that was used to represent the primary outcome "mental health" is listed. With BDI: Beck Depression Inventory; PHQ-8: Personal Health Questionnaire Depression scale; POMS: Profile of Mood States; GIDS-c: Quick Inventory of Depressive Symptoms-Clinician rated; MSP: Mesure du Stress Psychologique; LSAS-SR: Liebowitz Social Anxiety Scale Self-Report; MADRS: Montgomery–Åsberg Depression Rating Scale; SIAS: Social Interaction Anxiety Scale; BAI: Beck Anxiety Inventory; STAI: State-Trait Anxiety Inventory; DASS: Depression Anxiety Stress Scales; GSI: General Symptom Index; FQ: Fear Questionnaire; GAD7: Generalized Anxiety Disorder 7-item; BSI: Brief Symptom Inventory.

[e]Study is considered an outlier in within-subject analyses.

[f]The data used for the analyses consist of all pooled participants, the outcome questionnaire at pre-intervention is compared with last outcome questionnaire that participant completed.

[g]The intervention could be accessed using the mobile phone, tablet, and computer.

[h]Study is labeled as a prepost design because it is unclear whether participants were randomized across conditions.

[i]The study technically is an ecological momentary assessment study with an experimental manipulation.

Table 2. Characteristics of the ecological momentary intervention studies (part 2).

Study[a]	Intervention technique	Training type (+ type of MHP[b] support[c])	Training trigger	No. of training sessions[d]	Control (n)[e]
Included in meta-analysis					
Agyapong et al, 2012[f]	Self-management and monitoring	Passive (stand-alone + CAU)	Triggered	168 (2)	Waitlist (n=28)
Ahtinen et al, 2013	Acceptance and commitment therapy	Active	On-demand		
Aikens et al, 2015[g] (all pooled subjects)	Self-management and monitoring	Passive (+MHP)	Triggered	26 (1)	
Askins et al, 2009	Self-management and monitoring	Active (+MHP)	...	...	
Ben-Zeev et al, 2014	Self-management and monitoring	Active (+stand-alone + CAU)	Triggered	90 (3)	
Burns et al, 2011[f]	Behavioral activation	Active (+MHP)	Triggered	280 (5)	
Carissoli et al, 2015	Mindfulness	Active	On-demand	36 (2)	Placebo (n=18)
Dagöö et al, 2014[h] (mCBT[b])	Cognitive behavioral therapy	Active (+MHP)	...	...	
Dagöö et al 2014[h] (mIPT[b])	Interpersonal therapy	Active (+MHP)	...	...	
Depp et al, 2015	Self-management and monitoring	Passive (+MHP)	Triggered	140 (2)	Paper and pencil version (n=41)
Enock et al, 2014	Cognitive bias modification	Active	Triggered	84 (3)	Placebo (n=104)
Granholm et al, 2012	Cognitive behavioral therapy	Active (stand-alone + CAU)	Triggered	216 (3)	
Grassi et al, 2007 (Vnar[b])	Relaxation	Active	...	4 (2)	Waitlist (n=30)
Grassi et al, 2007 (Nnar[b])	Relaxation	Active	...	4 (2)	
Grassi et al, 2007[f] (MP3[b])	Relaxation	Active	...	4 (2)	
Harrison et al, 2011	Self-management and monitoring	Passive	On-demand	...	
Huffziger et al, 2013[i]	Mindfulness	Passive	Triggered	10 (10)	
Kenardy et al, 2003[f]	Cognitive behavioral therapy	Active (+MHP)	Triggered	420 (5)	CBT6 (n=44)
Lappalainen et al, 2013	Cognitive behavioral therapy and acceptance and commitment therapy	Active (+MHP)	On-demand	...	Waitlist (n=12)
Ly et al, 2014[f] behavioral activation	Behavioral activation	Active (+MHP)	...	...	
Ly et al, 2014 mindfulness	Mindfulness	Active (+MHP)	...	...	
Ly et al, 2014 mindfulness	Acceptance and commitment therapy	Active	On-demand	...	
Newman et al, 2014	Cognitive behavioral therapy	Active (+MHP)	Triggered	112 (4)	CBT6 (n=14)
Newman et al, 1997	Cognitive behavioral therapy	Active (+MHP)	Triggered	336 (4)	CBT12 (n=9)
Pallavicini et al, 2009 (VRMB[b])	Relaxation	Active (+MHP)	On-demand	...	Waitlist (n=4)

Study[a]	Intervention technique	Training type (+ type of MHP[b] support[c])	Training trigger	No. of training sessions[d]	Control (n)[e]
Pallavicini et al, 2009 (VRM[b])	Relaxation	Active (+MHP)	On-demand	...	
Proudfoot et al, 2013	Self-management and monitoring	Passive	On-demand	...	Placebo (n=195)
Repetto et al, 2013 (VRMB)	Relaxation	Active (+MHP)	On-demand	...	Waitlist (n=8)
Repetto et al, 2013 (VRM)	Relaxation	Active (+MHP)	On-demand	...	
Rizvi et al, 2011	Dialectical behavior therapy	Active (+TAU)	On-demand	...	
Shapiro et al, 2010	Self-management and monitoring	Passive (+MHP)	—	168 (1)	
Watts et al, 2013[f]	Cognitive behavioral therapy	Active (+MHP)	On-demand	...	Computer version (n=15)
Wenze et al, 2014	Cognitive behavioral therapy	Passive (stand-alone + CAU	Triggered	28 (2)	
Not included in meta-analysis					
Gorini et al, 2010 (VRMB)	Relaxation	Active (+MHP)	On-demand	...	Waitlist (n=8)
Gorini et al, 2010 (VRM)	Relaxation	Active (+MHP)	On-demand	...	
Grassi et al, 2011 (Vnar)	Relaxation	Active	...	6 (1)	Waitlist (n=15)
Grassi et al, 2011 (MP3[b])	Relaxation	Active	...	6 (1)	
Preziosa et al, 2009 (Vnar; study 1)	Relaxation	Active	...	6 (1)	Waitlist (n=6)
Preziosa et al, 2009 (MP3; study 1)	Relaxation	Active	...	6 (1)	
Riva et al, 2006	Relaxation	Active	...	4 (2)	Placebo (n=30)
Preziosa et al, 2009 (study 2)	Relaxation	Active	...	4 (2)	Placebo (n=11)
Zautra et al, 2012 (mindfulness)	Mindfulness	Active	Triggered	27 (1)	Placebo (n=23)
Zautra et al, 2012 (mastery-control)	Behavioral activation	Active	Triggered	27 (1)	

[a]Studies are ordered by inclusion in the meta-analysis. Behind the study's year of publication, between brackets, the sample (or condition) that received the EMI was specified.

[b]mCBT: mobile cognitive behavioral therapy; mIPT: mobile interpersonal psychotherapy; MP3: audio only condition; MHP: mental health professional; Nnar: video only condition; Vnar: video narrative condition; VRMB: virtual reality and mobile condition with biofeedback; VRM: virtual reality with mobile condition.

[c]Following the type of training, the type of support by the mental health professional is reported between brackets. With +MHP=mental health professional–supported EMI; stand-alone + CAU=stand-alone EMI with access to care as usual. No information was displayed when the EMI was stand-alone.

[d]The maximum number of total training sessions is reported. The maximum number of daily training sessions is reported between brackets.

[e]Control condition (and sample size at post-intervention) is listed if the study was included in the between-subject analyses. If the control condition is an active treatment, it is specified which specific active treatment condition is used to calculate the effect size. With CBT6=6-sessions of cognitive behavioral therapy; CBT12=12-sessions of cognitive behavioral therapy.

[f] Study is considered an outlier in within-subject analyses.

[g]The data used for the analyses consist of all pooled participants, the outcome questionnaire at preintervention is compared with last outcome questionnaire that participant completed.

[h]The intervention could be accessed using the mobile phone, tablet, and computer.

[i]The study is technically an ecological momentary assessment study with an experimental manipulation.

Figure 1. PRISMA flow diagram for study inclusion.

PRISMA 2009 Flow Diagram
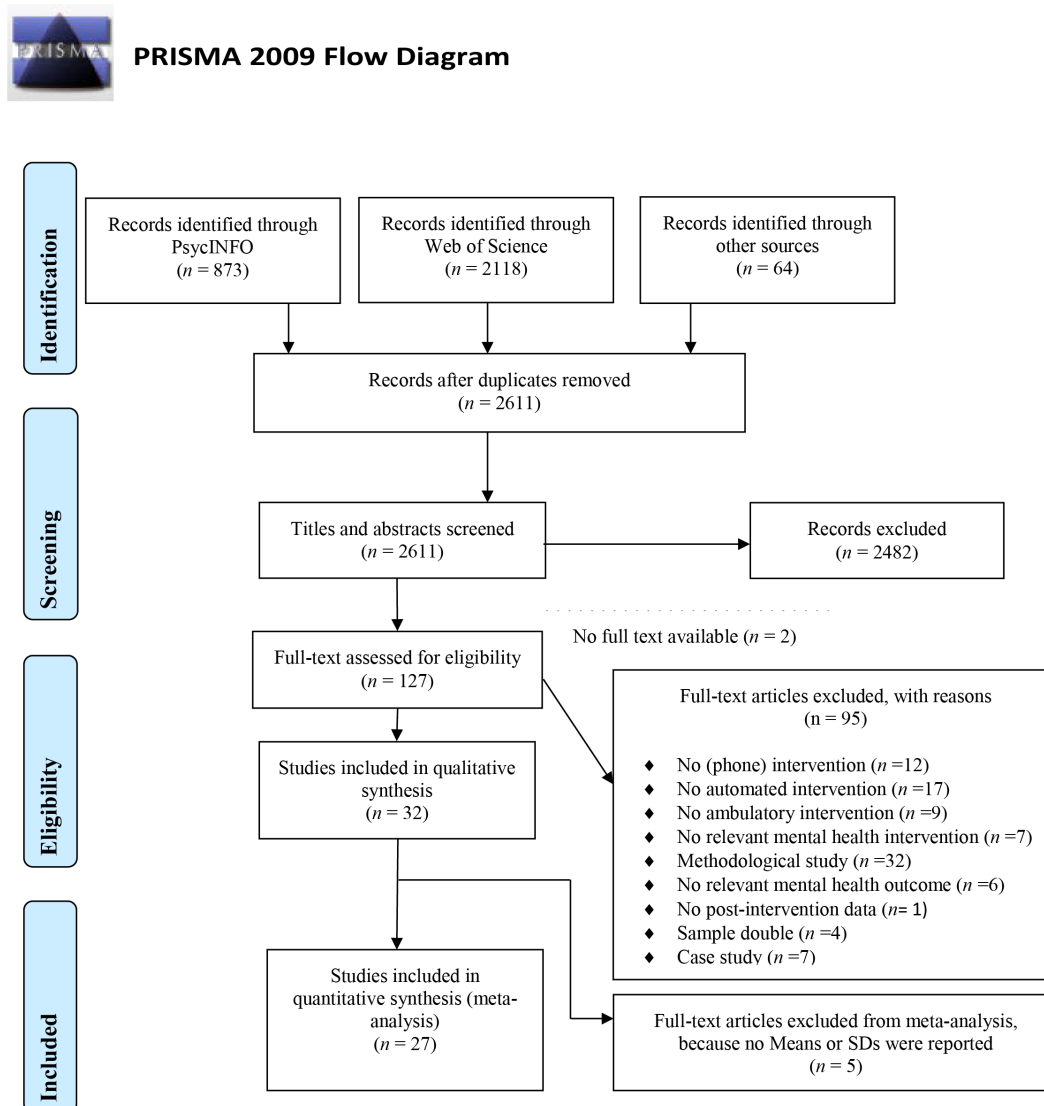
Study Characteristics

Of the 33 studies that were included, 17 had a prepost design, and 16 studies were an RCT. Of the total number of studies, 10 included healthy individuals [25,26,33,44,48-51] (studies 1 and 2 [45]), and the remaining studies focused on a clinical sample. Specifically, the focus of 8 studies was on anxiety disorders [31,32,35,36,43,52-54], 6 on depressive symptoms (ranging from mild symptoms to major depressive disorder) [34,47,55-58], 1 on perceived stress [59], 2 on anxiety, depression, and stress [60,61], 2 on bipolar disorder [62,63], 2 on schizophrenia [50,64], 1 on borderline personality disorder [65], and 1 on bulimia nervosa [66]. No study had positive psychological well-being as primary outcome. Across the studies, the average age ranged from 20.86 to 54.05 years with a mean of 37.33 (SD=9.37). Only female participants were included in 4 studies [44,48,66] (study 1 [45]), and 1 study included only males [59], and overall, the percentage of females was 64.79 (SD=22.72).

Intervention Characteristics

A range of different intervention techniques were studied: CBT [35,36,50,52,54,58,59,63], acceptance and commitment therapy [33,51,59], mindfulness [26,47,49,57], behavioral activation [47,56,57], relaxation [25,31,32,43-46], interpersonal therapy [52], dialectical behavior therapy [65], cognitive bias modification [53], and self-management and/or monitoring strategies [34,48,55,60-62,64,66]. The EMI was offered in combination with therapy in 10 studies (30%). Four studies combined the EMI with CBT [35,36,54,66], 3 with virtual reality including both relaxation and exposure [31,32,43], 1 with a problem-skill training [48], 1 with psychoeducation [62], and one with meetings including mindfulness and acceptance exercises [59]. In 5 studies, the EMI was a stand-alone intervention in combination with care as usual. This care focused on bipolar disorder [63], schizophrenia or schizoaffective disorder [50,64], major depressive disorder, and alcohol dependency [55], or on borderline personality disorder and substance abuse [65]. The other 18 studies investigated whether the use of an individual EMI can be effective without face-to-face therapy confounding the effect. Nevertheless, support by an MHP was included in 5 of these 18 studies. The MHP was for instance used to support the participant in the first phase of the intervention [58], to give feedback on the homework using Internet or email [52,57] or to increase adherence by telephone [34,56]. As can be seen in Table 2, 13

studies (39%) did not include support by an MHP after starting the EMI. In addition to the EMI and the potential support offered by the MHP, 6 of the 33 studies used a website for psychoeducation [51,57] or for providing therapy modules [56,59-61]. Most of the EMIs under investigation were "active" (25/33, 76%), meaning that participants had to carry out an exercise as part of the intervention. The EMIs in the remaining studies were classified as passive and only provided the participant with information.

On average, the EMI lasted for 7.47 weeks (SD=6.46), but this varied considerably. For example, the studies with the shortest EMI lasted only 1 or 2 days [25,26,46] (study 2 [45]), whereas the study with the longest EMI lasted for 26 weeks [34]. However, these numbers may be only modestly informative considering that the number of training episodes that people received (per day) varied highly across the studies. To explain, the study with the shortest length of training actually had the highest number of training episodes per day [26], whereas the study with the longest training length only trained people once a week [34]. Therefore, it may be more valuable to examine how many training episodes participants received per day and in total. Unfortunately, 13 studies did not specify the number of training episodes (per day or in total). Across the 20 other studies, the average number of training episodes was 2.80 per day (SD=2.12) ranging from 1 to 10, and on average 108.25 in total (SD=123.00) ranging from 4 to 420. The number of training episodes not only varied across studies but likely also varied across individuals within a given study. Fifteen of the 33 studies (ie, 45%) reported (some) information about compliance with the training, but the information used to represent compliance differed across studies. The average compliance with the sessions or treatment modules was 73.88% (SD=16.73) [26,47,50,52,53,57,58,60,62,63,66]. Burns et al [56] reported that the number of training sessions was on average 15.30 (SD=8.30) in the first week and that this decreased to 9.00 (SD=6.50) in the final week. In study of Ben-Zeev et al [64], participants used the training on 86.50% of the days and on these days used on average 5.19 sessions. Participants in the study by Aikens et al [34] participated in a median of 25 weeks (of the 26 weeks). Finally, Lappalainen et al [59] discloses that all participants tried at least 3 of the 6 available tools; however, no data are reported on the frequency of use.

The training episodes were automatically triggered by the device in 13 studies, and in 11 studies, the training episodes were not specifically triggered, and participants could complete the training whenever they wanted. Nine studies did not report whether the training was triggered or whether it was accessed on-demand.

Quality Assessment

The quality assessment of the studies is summarized in Table 3 and is on average 2.29 (SD=1.42, NB on a scale from 0 to 6), which can be considered low. Nine studies had a pre-intervention to post-intervention design, so the quality domain "selection bias"—as indexed by "random sequence generation" and "allocation concealment"—was not applicable (quality domain 1, see the previous section) [33,50,51,56,60,63-66]. Only 5 studies had a low risk of bias on this domain [52,57,58,61,62], with 5 other studies having a low risk of bias on "random sequence generation" and an unclear or high risk on "allocation concealment" [26,31,32,48,55]. In the remaining 14 studies, the risk was either unclear or high. The blinding of personnel (domain 2) was achieved in only 2 studies [61,62]. Moreover, most studies used self-report questionnaires, with only 2 studies using clinician-rated interviews (domain 3)—however, clinicians were not blinded for the condition of the participants [56,63]. There was a high risk for attrition (domain 4; ie, ≥ 20%) in 8 studies [48,50,53,58,60-62,66], and attrition (in the EMI group) was not disclosed in 7 studies [25,35,43,44,46] (studies 1 and 2 [45]). Finally, 7 studies failed to report the results for all prespecified outcome types (domain 5) [25,32,43,44,46] (studies 1 and 2 [45]).

Table 3. Quality assessment of the individual studies using the Cochrane Collaboration's tool.

Study	Random sequence generation[a]	Allocation concealment[a]	Performance bias[b]	Detection bias	Attrition bias[c]	Reporting bias[d]	Overall grade[e]
Agyapong et al, 2012	+	−	−	−	+	+	3
Ahtinen et al, 2013	N/A	N/A	−	−	+	+	4
Aikens et al, 2015	−	−	−	−	+	+	2
Askins et al, 2009	+	?	−	−	−	+	2
Ben-Zeev et al, 2014	N/A	N/A	−	−	+	+	4
Burns et al, 2011	N/A	N/A	−	?	+	+	4
Carissoli et al, 2015	?	?	−	−	+	+	2
Dagöö et al, 2014	+	+	−	−	+	+	4
Depp et al, 2015	+	+	+	−	−	+	4
Enock et al, 2014	?	?	?	−	−	+	1
Gorini et al, 2010[f]	?	?	−	−	?	−	0
Granholm et al, 2012	N/A	N/A	−	−	−	+	3
Grassi et al, 2011[f]	?	?	−	−	?	−	0
Grassi et al, 2007	?	?	−	−	?	−	0
Harrison et al, 2011	N/A	N/A	−	−	−	+	3
Huffziger et al, 2013	+	?	−	−	+	+	3
Kenardy et al, 2003	?	?	−	−	?	+	1
Lappalainen et al, 2013	?	?	−	−	+	+	2
Ly et al, 2014	+	+	−	−	+	+	4
Ly et al, 2012	N/A	N/A	−	−	+	+	4
Newman et al, 2014	?	?	−	−	+	+	2
Newman et al, 1997	?	?	−	−	+	+	2
Pallavicini et al, 2009	+	?	−	−	+	−	2
Preziosa et al, 2009[f] (studies 1 and 2)	?	?	−	−	?	−	0
Proudfoot et al, 2013	+	+	+	−	−	+	4
Repetto et al, 2013	+	?	−	−	+	−	2
Riva et al, 2006[f]	?	?	−	−	?	−	0
Rizvi et al, 2011	N/A	N/A	−	−	+	+	4
Shapiro et al, 2010	N/A	N/A	−	−	−	+	3
Watts et al. 2013	+	+	−	−	−	+	3
Wenze et al, 2014	N/A	N/A	−	?	+	+	4
Zautra et al, 2012[f]	?	?	−	−	+	+	2

[a]The label "not applicable" (N/A) is used in 1-armed studies.

[b]The risk for performance bias is rated low if personnel are blinded irrespective of whether participants were blinded.

[c]The bias for attrition is considered high when the attrition from pre-intervention to post-intervention is 20% or more.

[d]The bias for selective reporting is labeled low if all prespecified outcomes are reported, it is not necessary that all statistical information is reported per outcome (eg, means, standard deviation, CI, *P* values).

[e]The overall grade is determined by summing the number of low-risk categories and the number of N/A categories; +=low risk of bias; −=high risk of bias; ?=unclear risk of bias.

[f]Study is not included in the meta-analysis.

XSL•FO
RenderX

Within-Subject Analyses

A total of 27 publications including 33 EMI groups (n=1156), were included in the within-subject analyses, and these studies had significant heterogeneity, Q (32)=188.80 with P<.001. The I^2 statistic showed that the observed variance was high (I^2=83.05). This further supports the use of a random effect model in the analyses.

The average effect on mental health from pre-intervention to post-intervention was g=0.73, 95% CI (0.56-0.90), P<.001 (see Figure 2 and Table 4), indicating a medium to large effect. To determine whether there was a risk for publication bias, the distribution in the funnel plot was examined. As can be seen in Figure 3, most of the studies (white circles) are centered at the top of the plot and are distributed to the right side of the mean as the sample size decreases. This reflects the presence of a publication bias, and an Egger's test of intercept was used as a method to quantify the amount of bias. In this case, the intercept was 1.89, 95% CI (0.28-3.51), with t (31)=2.392 and 1-sided P=.01. In other words, there was a significant risk for bias. To correct for the missing studies to the left of the mean, the trim and fill method was used. Figure 3 shows that 2 studies (black circles) were added and the corrected effect size was g=0.70, 95% CI (0.52-0.87). The corrected effect is virtually identical to the unadjusted effect, which suggests that the reported findings are quite robust and are not simply due to publication bias.

The standardized residual identified 6 studies as outliers, and these were removed from the analyses [35,55,56,58] (MP3 condition [25]) (BA condition [57]). Removal of these studies resulted in a decrease in effect and heterogeneity (g=0.57, 95% CI: 0.45-0.70, P<.001; Q (26)=74.46, I^2=65.08). Nevertheless, the effect was still medium for the 27 included EMI groups (n=1008), and the studies were significantly heterogeneous.

It was explored whether the effect was different per outcome type. Depressive symptoms were assessed in 17 studies; anxiety in 15 studies; quality of life in 6 studies; stress in 5 studies; acceptance in 4 studies, and relaxation in 3 studies. As can be seen in Table 5, there was evidence for an effect on anxiety (g=0.47, 95% CI: 0.32-0.63, P<.001), depression (g=0.48, 95% CI: 0.34-0.61, P<.001), perceived stress (g=0.40, 95% CI: 0.23-0.57, P<.001), acceptance (g=0.36, 95% CI: 0.13-0.59, P=.002), and quality of life (g=0.38, 95% CI: 0.19-0.56, P<.001). No effect was found on relaxation with g=0.28, 95% CI (−0.46 to 1.01), P=.46. However, there was no evidence that the effect differed significantly per outcome type with Q (5)=1.74, P=.88.

Furthermore, subgroup analyses were done to see whether the effect varied by moderator. Table 4 shows that "support by an MHP" was the only moderator for which the effect varied significantly, Q (2)=6.77, P=.03. Specifically, the effect was medium to large when the EMI included support by an MHP (g=0.73, 95% CI: 0.57-0.88), small to medium for the stand-alone EMI (g=0.45, 95% CI: 0.22-0.69), and small for those individuals who received a stand-alone EMI in combination with care as usual (g=0.38, 95% CI: 0.11-0.64).

Table 4. Effect sizes (Hedges' g) of ecological momentary intervention on mental health by study and intervention characteristics (within-subject analyses)[a].

Outcome			Random effect model			Heterogeneity		Test of difference
			k [b]	n[c]	g (95% CI)[d]	Q [e]	I [e]	Q [f]
Mental health			27	1008	0.57 (0.45-0.70)[g]	74.46[g]	65.08	
	Design							1.03
		RCT[h]	11	481	0.65 (0.48-0.82)[g]	24.10[i]	58.50	
		Pre-post	16	527	0.52 (0.33-0.71)[g]	47.34[g]	68.32	
	Sample							1.79
		Clinical	20	793	0.63 (0.50-0.76)[g]	39.32[i]	51.68	
		Healthy	7	215	0.40 (0.10-0.71)[j]	26.76[g]	77.58	
	Age[k], years							2.19
		≤ 38.15	12	426	0.61 (0.36-0.86)[g]	54.38[g]	79.77	
		> 38.15	12	552	0.51 (0.37-0.64)[g]	17.64[l]	37.65	
		Unspecified	3	30	0.80 (0.41-1.18)[g]	0.40	0.00	
	Gender[k]						1.96	
		≤ 60% female	14	450	0.49 (0.28-0.70)[g]	51.25[g]	74.63	
		> 60% female	11	550	0.67 (0.53-0.81)[g]	15.94	37.26	
		Unspecified	2	8	0.55 (–0.08 to 1.17)[l]	1.12	10.43	
	Sample size[k]							1.18
		≤ 22 participants	13	158	0.67 (0.46-0.87)[g]	17.24	30.39	
		> 22 participants	14	850	0.52 (0.36-0.69)[g]	56.36[g]	76.93	
	Training type							0.32
		Active	20	518	0.60 (0.42-0.78)[g]	57.51[g]	66.96	
		Passive	7	490	0.53 (0.34-0.71)[g]	16.65[j]	63.97	
	Training trigger							1.65
		Triggered	9	535	0.52 (0.33-0.71)[g]	26.96[i]	70.45	
		On-demand	11	256	0.49 (0.37-0.62)[g]	9.41	0.00	
		Unspecified	7	217	0.76 (0.38-1.14)[g]	35.69[g]	83.19	
	No. of daily training episodes[k]							0.53
		≤ 2	7	370	0.55 (0.24-0.87)[i]	32.65[g]	81.62	
		> 2	6	259	0.51 (0.20-0.82)[i]	22.81[g]	78.08	
		Unspecified	14	379	0.63 (0.49-0.77)[g]	17.48	25.62	
	No. of total training episodes[k]							0.92
		≤ 84	7	481	0.48 (0.21-0.75)[i]	36.62[g]	83.62	
		> 84	6	148	0.62 (0.27-0.97)[i]	17.77[i]	71.86	
		Unspecified	14	379	0.63 (0.49-0.77)[g]	17.48	25.62	

XSL•FO
RenderX

Outcome	Random effect model			Heterogeneity		Test of difference
	k [b]	n[c]	g (95% CI)[d]	Q [e]	I [e]	Q [f]
Support MHP[m]						6.77[j]
MHP-supported EMI	14	474	0.73 (0.57-0.88)[g]	20.67[l]	37.10	
Stand-alone EMI	9	425	0.45 (0.22-0.69)[g]	35.81[j]	77.66	
Stand-alone EMI with access to care as usual	4	109	0.38 (0.11-0.64)[i]	5.37	43.97	
Quality assessment[k]						0.01
≤ 3	17	781	0.57 (0.39-0.76)[g]	57.68[j]	72.26	
> 3	10	227	0.59 (0.42-0.76)[g]	16.78[l]	46.38	

[a]Outliers were excluded from the presented moderation analyses (ie, 6 studies).

[b]k=number of studies.

[c]n=number of participants.

[d]g=effect size Hedges' g with 95% CI.

[e]Q and I^2=heterogeneity statistics.

[f]Q=contrast between subgroups.

[g]$P<.001$.

[h]RCT=randomized controlled trial.

[i]$P<.01$.

[j]$P<.05$.

[k]Data were categorized based on the median.

[l]$P<.10$.

[m]MHP=mental health professional.

Table 5. Effect sizes (Hedges' g) of ecological momentary intervention by outcome type (within-subject analyses)[a].

Outcome	Random effect model			Heterogeneity		Test of difference
	k [b]	n[c]	g (95% CI)[d]	Q [e]	I [e]	Q [f]
Overall	50	1830				1.74
Anxiety	15	468	0.47 (0.32-0.63)[g]	28.28[h]	50.49	
Depression	17	870	0.48 (0.34-0.61)[g]	46.48[g]	65.58	
Perceived stress	5	199	0.40 (0.23-0.57)[g]	4.59	12.79	
Relaxation	3	106	0.28 (−0.46 to 1.01)	25.28[g]	92.09	
Acceptance	4	72	0.36 (0.13-0.59)[i]	2.79	0.00	
Quality of life	6	115	0.38 (0.19-0.56)[g]	4.25	0.00	

[a]Outliers were excluded from the presented moderation analyses (ie, 6 studies).

[b]k=number of studies.

[c]n=number of participants.

[d]g=effect size Hedges' g with 95% confidence interval.

[e]Q and I^2=heterogeneity statistics.

[f]Q=contrast between subgroups.

[g]$P<.001$.

[h]$P<.05$.

[i]$P<.01$.

XSL•FO
RenderX

Figure 2. Forest plot showing the effect of ecological momentary interventions (EMIs) on mental health complaints for all within-subject studies. The EMI sample (or condition) is reported after the year of publication when multiple EMI samples were included in a publication.
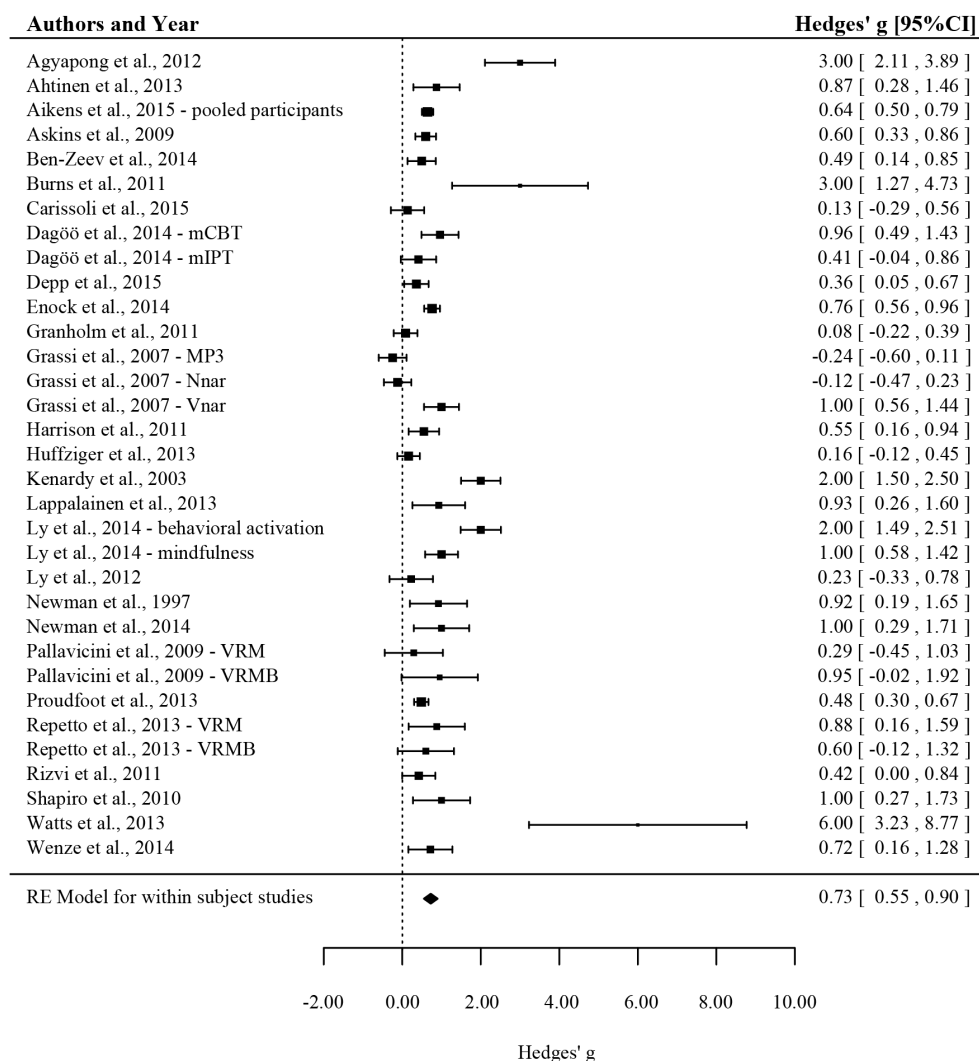
Figure 3. Funnel plot of standard error by Hedges' *g* with imputed values based on Duval and Tweedie's trim and fill method (within-subject studies).
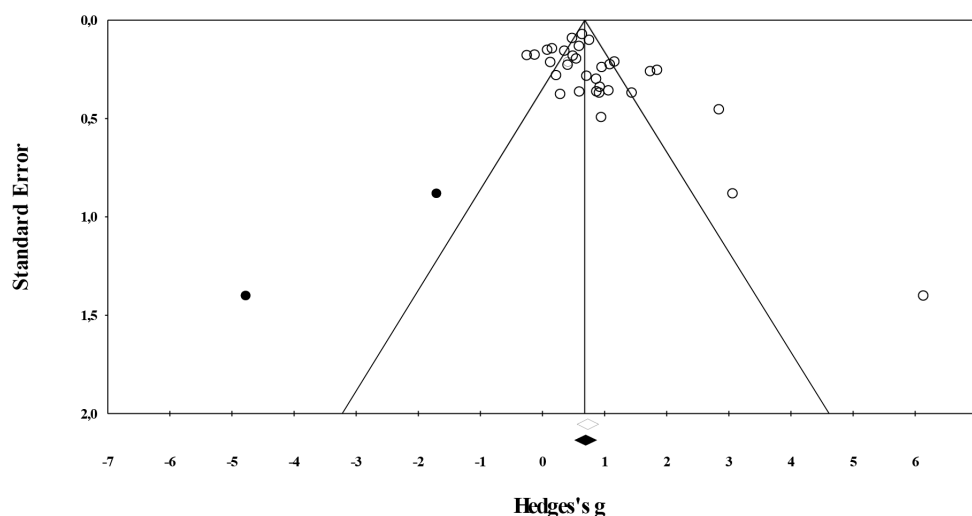
Between-Subject Analyses

In the between-subject analyses, only 1 EMI group per study was included (see "Coding"). A total of 13 studies were included with 454 participants in the EMI condition and 522 participants in a control condition (waitlist, placebo, or active treatment control). The included studies were not significantly heterogeneous, Q (12)=17.17, P=.14. Moreover, the observed true variance was small (I^2=30.13). A small value of I^2 indicates that a large part of the variance is the result of random error. If one tries to explain this variance (with subgroup analyses), one tries to find an explanation for something that is in essence random [30]. Therefore, no attempt will be made to explain the variance in effect by testing differences due to outcome types and other moderators. Still, a random effect model was adopted because we do not assume a common effect size (despite the lack of statistical significant variance between studies) [30].

The effect for EMI in between-subject studies was g=0.40, 95% CI (0.22-0.57), P<.001 (see Figure 4). This effect can be considered small to medium. The funnel plot (see Figure 5) shows that there is indication for publication bias; the distribution of effects is asymmetrical as the sample size decreases. Specifically, effect sizes are more likely to fall to the right side of the mean when the sample size is small. Furthermore, the Egger's test of intercept is significant, indicating that there is a risk for bias (intercept is 1.50, 95% CI: 0.28-2.72) with t (11)=2.708, 1-sided P=.01). The trim and fill method was used to account for the missing studies. Six studies were added to the left of the mean (black circles in Figure 5), and the corrected effect size was g=0.23, 95% CI (0.04-0.42). The corrected effect is considerably smaller than the uncorrected effect, which indicates that the uncorrected effect may be subject to publication bias and needs to be interpreted carefully. On the basis of the standardized residuals, no study was identified as an outlier.

Figure 4. Forest plot showing the effect of ecological momentary interventions (EMIs) on mental health complaints for all between-subject studies. The EMI sample (or condition) that was used to represent the active treatment condition is reported after the year of publication.

Authors and Year	Hedges' g [95% CI]
Agyapong et al., 2012	0.83 [0.27 , 1.39]
Carissoli et al., 2015	0.11 [-0.52 , 0.73]
Depp et al., 2015	0.47 [0.03 , 0.90]
Enock et al., 2014	0.04 [-0.22 , 0.31]
Grassi et al., 2007 - Vnar	0.75 [0.23 , 1.27]
Kenardy et al., 2003	0.24 [-0.18 , 0.67]
Lappalainen et al., 2013	0.60 [-0.21 , 1.41]
Newman et al., 1997	0.38 [-0.50 , 1.27]
Newman et al., 2014	0.66 [-0.13 , 1.44]
Pallavicini et al., 2009 - VRMB	0.64 [-0.60 , 1.89]
Proudfoot et al., 2013	0.27 [0.04 , 0.49]
Repetto et al., 2013 - VRMB	0.46 [-0.50 , 1.43]
Watts et al., 2013	1.20 [0.36 , 2.05]
RE Model for between-subject studies	0.40 [0.22 , 0.57]

Hedges' g

Figure 5. Funnel plot of standard error by Hedges' *g* with imputed values based on Duval and Tweedie's trim and fill method (between-subject studies).
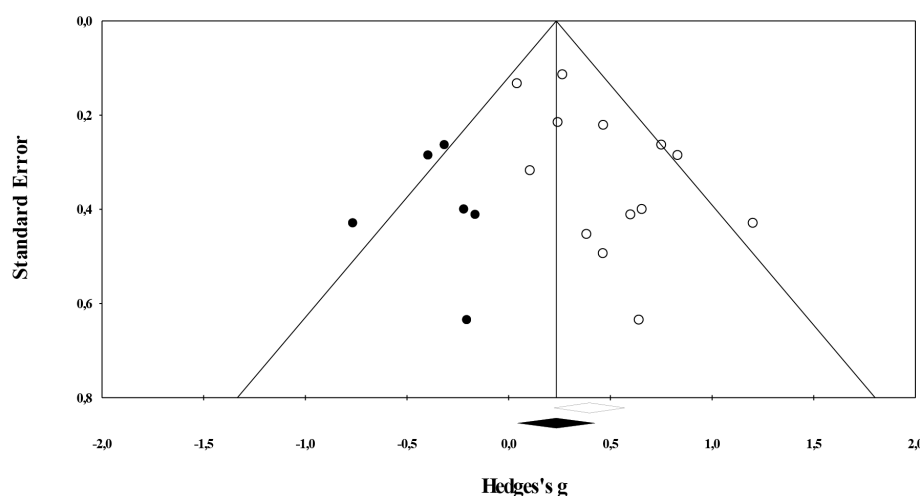
Discussion

Principal Findings

The systematic review and meta-analysis was a first attempt to examine whether mobile technologies can be used to provide an effective intervention for mental health and under which circumstances this is the case. A total of 33 studies (*n*= 1301) were used to answer this question, and the included studies varied considerably in terms of study and intervention characteristics. The quality assessment indicated that the reported study quality was generally low. Specifically, the studies were at risk for bias caused by attrition, reliance on self-report measures, and the failure to blind personnel. Moreover, only a few studies reported using strategies to randomly allocate participants to conditions.

In the within-subject studies (n=1008), a significant medium effect size (Hedges' *g*) of 0.58 was found. The estimated effect size did not significantly differ per outcome type (ie, anxiety, depression, perceived stress, acceptance, relaxation, and quality of life), although no significant effect was found for relaxation. Moderation analysis suggested that the effect on mental health was 62% larger when the EMI was part of a treatment package that included support of an MHP compared with stand-alone EMI. Moreover, this moderation analyses showed that the effect of EMI was smaller, but significant, in the population that had access to care as usual while using the EMI (eg, inpatient or outpatient setting). It is possible to speculate about what caused this difference in effect; however, a clear comparison of the groups is complicated by the fact that the groups (and included studies) are very diverse. More specifically, the group that received EMIs while also having access to care as usual consisted largely of patients with severe complaints that might be less susceptible to change (eg, schizophrenia or schizoaffective disorders, borderline personality disorder, and substance abuse).

With regard to the between-subject studies (n=454), the estimated effect size was 0.40. The effect was, however, subject to publication bias, and the corrected effect was considered small, but significant (*g*=0.23).

Both the within- and the between-subject analyses indicate that mobile technologies can be effectively used to deliver interventions for mental health. When interpreting this effect, it must be acknowledged that the effects were considerable smaller in the between-subject studies compared with the within-subject studies. A larger effect in within-subject studies is frequently observed. However, within-subject studies are limited because causality can—generally—not be interfered from these studies. Moreover, these studies have an increased risk for type-II errors, which implies that the conclusions from within-subject studies must be interpreted with caution [67]. Nevertheless, both study types provide a first—and positive—insight into how mobile technology can be used to improve mental health.

The finding that the effect of EMIs was stronger when support by an MHP was included is in line with findings from research on Internet interventions (eg, [68,69]). Therefore, although fully automated EMIs can have a positive effect on mental health, it is additionally beneficial to include contact between researcher (or therapist) and participant. This contact could be a helpful tool to increase adherence and motivation, which in turn could result in a stronger effect. Unfortunately, it is currently unknown what levels of support are needed to optimize the effectiveness of EMIs. Future studies should differentiate what kind of contact is necessary for improvement. Not only is it important that we learn how much contact is required, but the *when* (eg, beginning or during intervention), *how* (eg, via mobile phone, email, or face-to-face), and *what* (eg, should support focus on adherence or on the intervention) questions are also worth asking when developing evidence-based interventions [69]. In addition, it is worthwhile to consider which individuals stand to benefit from the support and if support is necessary for everyone. To specify, EMIs can be a valuable (first) step to treat the "worried well" and individuals with mild symptoms. Using EMIs to treat this group could be economically efficient, as mild problems constitute a major part of all reported mental health problems [70]. Treating this group using the cost-effective EMI

methodology, frees resources (such as therapists) for those individuals who are in greater need of more intensive interventions. Moreover, it could help to improve the access to and quality of psychological care. Ideally, the progress of the individuals using the EMIs could be monitored so that alternative intervention options can be recommended when an EMI fails to be effective. Alternative intervention options could entail extra support (while using the EMI), an Internet intervention, or face-to-face intervention. Incorporating EMI in a stepped-care program could help in providing intensive intervention only when needed [71].

Apart from the moderator "support by an MHP," no moderation effects were found for the other study or intervention characteristics. The intervention was, for example, equally effective for healthy versus clinical individuals. The absence of significant moderator variables implies that any form of EMI, irrespective of for instance type of training or number of training episodes, is equally effective for all individuals. Obviously, this assumption is implausible, and it is more likely that the null findings are the result of the relative small number of studies that specifically reported the intervention characteristics (eg, number of training episodes and whether training was triggered) [72]. Considering that the research field of EMIs is relatively new, it is understandable that limited information is available on what characteristics of an intervention are considered effective (or active). It does, however, highlight the need for research that determines what the active features of an intervention are [73]. Potential questions that could be targeted relate to the frequency and duration of the intervention (eg, is daily practicing required, and if so, how many times a day?). Although initial research suggests that (daily) repetition is necessary to learn a new behavior [74], this should be further investigated using RCTs with EMIs. Another potential research endeavor is whether a training should be offered on-demand or whether it should be automatically triggered. A meta-analysis, investigating the use of triggers to stimulate engagement with digital interventions, found preliminary support for the use of technology (eg, texting or emails) to improve engagement [75]. This result is interesting, as mobile interventions would make it easy to trigger a training, but more studies are needed to establish if this effect is valid. Altogether, it is important that future research focuses on identifying the most potent feature(s) of an intervention.

Limitations

This meta-analysis is limited by the low reported study quality (ie, 2.29 on a scale from 0 to 6). When the reported study quality is low, the study may be subject to weakness in the experimental setup or to problems in the processing of the data. These shortcomings can influence the true effect and lead to an overrepresentation or underrepresentation [38]. However, reported study quality must not be confused with the actual quality of the study. To explain, studies may have used excellent set-ups but may have failed to adequately report their precise procedure. Indeed, most of the studies failed—on one or more occasions—to provide sufficient information to establish whether there was a risk of bias. To perform correct quality assessments, it is recommended that authors of future studies follow publication guidelines such as the CONSORT statement for RCT [76].

In line with the previous limitation, it is also important that sufficient intervention details are described so that other researchers can fully comprehend what the intervention entailed. In the included studies, the content of the intervention was described, yet other important intervention components—as suggested by Davidson et al [28]—were not always disclosed. For instance, 10 of the 33 studies (30%) failed to report how the intervention was triggered, and more than half of the studies did not explicate what the compliance with the intervention was. It is imperative that studies describe the full details of used intervention and the compliance with the intervention, and the guidelines by Davidson et al [28] can be used for this purpose. This information can ultimately be used to determine which interventions (or intervention characteristics) are the most effective.

Another limitation is that the larger part of the included studies used a within-subject design. Although this design can yield valuable information, RCTs (which use a between-subject design) are considered superior when evaluating interventions because these can be used to establish a causal relation. Moreover, some of the included studies (both within- and between-subject) had small sample sizes. Studies with small sample sizes may be statistically underpowered to detect an effect and have a lower study validity [72,77]. To further strengthen the body of knowledge on the effectiveness of EMIs, RCTs using adequate numbers of participants are needed.

Conclusions

To conclude, the meta-analysis found a small to medium effect of EMIs on mental health, and this effect did not differ across the different outcome types. Furthermore, the effect appeared to be larger when the EMI was supported by an MHP. It is important that future research determines how support by an MHP can best be implemented and if this support is a necessity for everyone. In addition, new research studies should investigate what the active features of an EMI are. Overall, the use of EMIs for improving mental health is supported; EMIs offer great potential for providing easy and cost-effective strategies to improve mental health and positive psychological well-being in the population.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1

Specific search strings used to find publications.

[PDF File (Adobe PDF File), 6KB - jmir_v18i6e152_app1.pdf]

References

1. Kessler RC, Angermeyer M, Anthony JC, Demyttenaere K, Gasquet I, Gluzman S, et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. World Psychiatry 2007 Oct;6(3):168-176 [FREE Full text] [Medline: 18188442]

2. Alonso J, Angermeyer MC, Bernert S, Bruffaerts R, Brugha TS, Bryson H, ESEMeD/MHEDEA 2000 Investigators, European Study of the Epidemiology of Mental Disorders (ESEMeD) Project. Use of mental health services in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatr Scand Suppl 2004(420):47-54. [doi: 10.1111/j.1600-0047.2004.00330.x] [Medline: 15128387]

3. Bijl RV, de GR, Hiripi E, Kessler RC, Kohn R, Offord DR, et al. The prevalence of treated and untreated mental disorders in five countries. Health Aff (Millwood) 2003;22(3):122-133 [FREE Full text] [Medline: 12757277]

4. World Health Organization. 2013 May 27. Comprehensive mental health action plan URL: http://apps.who.int/gb/ebwha/pdf_files/WHA66/A66_R8-en.pdf?ua=1 [WebCite Cache ID 6f1aYPJcB]

5. Wanless D. Securing our Future Health: Taking a Long-Term View. URL: http://si.easp.es/derechosciudadania/wp-content/uploads/2009/10/4.Informe-Wanless.pdf [accessed 2016-02-03] [WebCite Cache ID 6f1b7iU7h]

6. Mehl R, Conner T. Handbook of research methods for studying daily life. New York: Guilford Press; 2012.

7. Proudfoot J, Parker G, Hadzi PD, Manicavasagar V, Adler E, Whitton A. Community attitudes to the appropriation of mobile phones for monitoring and managing depression, anxiety, and stress. J Med Internet Res 2010;12(5):e64 [FREE Full text] [doi: 10.2196/jmir.1475] [Medline: 21169174]

8. Olff M. Mobile mental health: a challenging research agenda. Eur J Psychotraumatol 2015;6:27882 [FREE Full text] [Medline: 25994025]

9. Price M, Yuen EK, Goetter EM, Herbert JD, Forman EM, Acierno R, et al. mHealth: a mechanism to deliver more accessible, more effective mental health care. Clin Psychol Psychother 2014;21(5):427-436 [FREE Full text] [doi: 10.1002/cpp.1855] [Medline: 23918764]

10. Schwabe L, Wolf OT. Stress prompts habit behavior in humans. J Neurosci 2009 Jun 3;29(22):7191-7198 [FREE Full text] [doi: 10.1523/JNEUROSCI.0979-09.2009] [Medline: 19494141]

11. Soares JM, Sampaio A, Ferreira LM, Santos NC, Marques F, Palha JA, et al. Stress-induced changes in human decision-making are reversible. Transl Psychiatry 2012;2:e131 [FREE Full text] [doi: 10.1038/tp.2012.59] [Medline: 22760555]

12. Schwabe L, Wolf OT. Socially evaluated cold pressor stress after instrumental learning favors habits over goal-directed action. Psychoneuroendocrinology 2010 Aug;35(7):977-986. [doi: 10.1016/j.psyneuen.2009.12.010] [Medline: 20071096]

13. Otto AR, Raio CM, Chiang A, Phelps EA, Daw ND. Working-memory capacity protects model-based learning from stress. Proc Natl Acad Sci U S A 2013 Dec 24;110(52):20941-20946 [FREE Full text] [doi: 10.1073/pnas.1312011110] [Medline: 24324166]

14. Kalichman S, Stein JA, Malow R, Averhart C, Dévieux J, Jennings T, et al. Predicting protected sexual behaviour using the Information-Motivation-Behaviour skills model among adolescent substance abusers in court-ordered treatment. Psychol Health Med 2002;7(3):327-338 [FREE Full text] [doi: 10.1080/13548500220139368] [Medline: 19079796]

15. Neal DT, Wood W, Quinn JM. Habits? A Repeat Performance. Current Directions in Psychol Sci 2006 Aug;15(4):198-202. [doi: 10.1111/j.1467-8721.2006.00435.x]

16. International Telecommuncation Union. 2005-2015 ICT data for the world, by geographic regions and by level of development URL: http://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2015/ITU_Key_2005-2015_ICT_data.xls [accessed 2016-02-03] [WebCite Cache ID 6f1dI1gYT]

17. Kazdin AE, Blase SL. Rebooting Psychotherapy Research and Practice to Reduce the Burden of Mental Illness. Perspect Psychol Sci 2011 Jan;6(1):21-37. [doi: 10.1177/1745691610393527] [Medline: 26162113]

18. Kaplan RM, Stone AA. Bringing the laboratory and clinic to the community: mobile technologies for health promotion and disease prevention. Annu Rev Psychol 2013;64:471-498. [doi: 10.1146/annurev-psych-113011-143736] [Medline: 22994919]

19. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39 [FREE Full text] [doi: 10.1348/135910709X466063] [Medline: 19646331]

XSL•FO
RenderX

20. Serino S, Triberti S, Villani D, Cipresso P, Gaggioli A, Riva G. Toward a validation of cyber-interventions for stress disorders based on stress inoculation training: a systematic review. Virtual Reality 2013 Oct 26;18(1):73-87. [doi: 10.1007/s10055-013-0237-6]

21. Marsch LA. Leveraging technology to enhance addiction treatment and recovery. J Addict Dis 2012;31(3):313-318 [FREE Full text] [doi: 10.1080/10550887.2012.694606] [Medline: 22873192]

22. 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]

23. Ehrenreich B, Righter B, Rocke DA, Dixon L, Himelhoch S. Are mobile phones and handheld computers being used to enhance delivery of psychiatric treatment? A systematic review. J Nerv Ment Dis 2011 Nov;199(11):886-891. [doi: 10.1097/NMD.0b013e3182349e90] [Medline: 22048142]

24. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. J Med Internet Res 2013;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]

25. Grassi A, Preziosa A, Villani D, Riva G. A relaxing journey: the use of mobile phones for well-being improvement. Annual Review of Cybertherapy and Telemedicine. 2007. p. 123-131 URL: https://fe083ca7-a-62cb3a1a-s-sites.googlegroups.com/site/unicatt-it-arctt-teams-edition-backup/volume-5--summer-2007/Grassi_ARCTT_2007.pdf?attachauth=ANoY7cqPfJtdFEuboV2VVsvqkJ-AobMxbhlRWZ3tG5pI8spMXlsfOO2LthN3vKefal8wYT-rk2jTdw54BDaQ8geXQdqxdJlmUTHOCE3LNfw8K2a3Q8VFiNAX3C5yUH3SwxitlwHoaPt3P2UuhG_m8C9iUbnh98X3oI3gF4p4j_1OxrN1FXg7QxaFaJccjYFQgsET6a7YF7jzScSldGubRaXvnnL4pboQW8U0LtZo5MovbDOBW8LOd_oTTavyazhV3fiHSl9RPqLlHc7WGsHDoE5NTbaP-9LXR-_ugfIE2t4xmY5NLeQqZbI%3D&attredirects=0 [WebCite Cache ID /6i7IjVV8j]

26. Huffziger S, Ebner-Priemer U, Eisenbach C, Koudela S, Reinhard I, Zamoscik V, et al. Induced ruminative and mindful attention in everyday life: an experimental ambulatory assessment study. J Behav Ther Exp Psychiatry 2013 Sep;44(3):322-328 [FREE Full text] [doi: 10.1016/j.jbtep.2013.01.007] [Medline: 23466521]

27. Huppert FA. Psychological well-beingvidence regarding its causes and consequences. Appl Psychol Health Well-Being Jul 2009 Mar 09;1(2):137-164. [doi: 10.1111/j.1758-0854.2009.01008.x]

28. Davidson KW, Goldstein M, Kaplan RM, Kaufmann PG, Knatterud GL, Orleans CT, et al. Evidence-based behavioral medicine: what is it and how do we achieve it? Ann Behav Med 2003 Dec;26(3):161-171. [Medline: 14644692]

29. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]

30. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to meta-analysis. Oxford: Wiley; 2009.

31. Repetto C, Gaggioli A, Pallavicini F, Cipresso P, Raspelli S, Riva G. Virtual reality and mobile phones in the treatment of generalized anxiety disorders: a phase-2 clinical trial. Pers Ubiquit Comput 2013 Feb;17(2):253-260. [doi: 10.1007/s00779-011-0467-0]

32. Pallavicini F, Algeri D, Repetto C, Gorini A, Riva G. Biofeedback virtual reality and mobile phones in the treatment of generalized anxiety disorder (GAD): a phase-2 controlled clinical trial. J Cyber Ther Rehabil. 2009. (4) p. 315-327 URL: https://www.researchgate.net/publication/289089225_Biofeedback_virtual_reality_and_mobile_phones_in_the_treatment_of_generalized_anxiety_disorder_gad_A_phase-2_controlled_clinical_trial [WebCite Cache ID 6i7ImMRbF]

33. Ahtinen A, Mattila E, Välkkynen P, Kaipainen K, Vanhala T, Ermes M, et al. Mobile mental wellness training for stress management: feasibility and design implications based on a one-month field study. JMIR Mhealth Uhealth 2013;1(2):e11 [FREE Full text] [doi: 10.2196/mhealth.2596] [Medline: 25100683]

34. Aikens JE, Trivedi R, Heapy A, Pfeiffer PN, Piette JD. Potential Impact of Incorporating a Patient-Selected Support Person into mHealth for Depression. J Gen Intern Med 2015 Jun;30(6):797-803. [doi: 10.1007/s11606-015-3208-7] [Medline: 25666218]

35. Kenardy JA, Dow MG, Johnston DW, Newman MG, Thomson A, Taylor CB. A comparison of delivery methods of cognitive-behavioral therapy for panic disorder: an international multicenter trial. J Consult Clin Psychol 2003 Dec;71(6):1068-1075. [doi: 10.1037/0022-006X.71.6.1068] [Medline: 14622082]

36. Newman MG, Przeworski A, Consoli AJ, Taylor CB. A randomized controlled trial of ecological momentary intervention plus brief group therapy for generalized anxiety disorder. Psychotherapy (Chic) 2014 Jun;51(2):198-206 [FREE Full text] [doi: 10.1037/a0032519] [Medline: 24059730]

37. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282 [FREE Full text] [Medline: 23092060]

38. Higgins Julian P T, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928 [FREE Full text] [Medline: 22008217]

39. Follmann D, Elliott P, Suh I, Cutler J. Variance imputation for overviews of clinical trials with continuous response. J Clin Epidemiol 1992 Jul;45(7):769-773. [Medline: 1619456]

40. Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [Medline: 19565683]

41. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003 Sep 6;327(7414):557-560 [FREE Full text] [doi: 10.1136/bmj.327.7414.557] [Medline: 12958120]

42. Viechtbauer W. Conducting meta-analysis in R with the metafor package. J Stat Softw Aug 2010;36(3):1-48. [doi: 10.18637/jss.v036.i03]

43. Gorini A, Pallavicini F, Algeri D, Repetto C, Gaggioli A, Riva G. Virtual reality in the treatment of generalized anxiety disorders. Stud Health Technol Inform 2010;154:39-43. [Medline: 20543266]

44. Grassi A, Gaggioli A, Riva G. New technologies to manage exam anxiety. Stud Health Technol Inform 2011;167:57-62. [Medline: 21685642]

45. Preziosa A, Grassi A, Gaggioli A, Riva G. Therapeutic applications of the mobile phone. British Journal of Guidance & Counselling 2009 Aug;37(3):313-325. [doi: 10.1080/03069880902957031]

46. Riva G, Preziosa A, Grassi A, Villani D. Stress management using UMTS cellular phones: a controlled trial. Stud Health Technol Inform 2006;119:461-463. [Medline: 16404099]

47. Zautra AJ, Davis MC, Reich JW, Sturgeon JA, Arewasikporn A, Tennen H. Phone-based interventions with automated mindfulness and mastery messages improve the daily functioning for depressed middle-aged community residents. Journal of Psychotherapy Integration 2012;22(3):206-228. [doi: 10.1037/a0029573]

48. Askins MA, Sahler OJ, Sherman SA, Fairclough DL, Butler RW, Katz ER, et al. Report from a multi-institutional randomized clinical trial examining computer-assisted problem-solving skills training for English- and Spanish-speaking mothers of children with newly diagnosed cancer. J Pediatr Psychol 2009 Jun;34(5):551-563 [FREE Full text] [doi: 10.1093/jpepsy/jsn124] [Medline: 19091804]

49. Carissoli C, Villani D, Riva G. Does a meditation protocol supported by a mobile application help people reduce stress? Suggestions from a controlled pragmatic trial. Cyberpsychol Behav Soc Netw 2015 Jan;18(1):46-53. [doi: 10.1089/cyber.2014.0062] [Medline: 25584730]

50. Granholm E, Ben-Zeev D, Link PC, Bradshaw KR, Holden JL. Mobile Assessment and Treatment for Schizophrenia (MATS): a pilot trial of an interactive text-messaging intervention for medication adherence, socialization, and auditory hallucinations. Schizophr Bull 2012 May;38(3):414-425 [FREE Full text] [doi: 10.1093/schbul/sbr155] [Medline: 22080492]

51. Ly KH, Dahl J, Carlbring P, Andersson G. Development and initial evaluation of a smartphone application based on acceptance and commitment therapy. Springerplus 2012;1:11 [FREE Full text] [doi: 10.1186/2193-1801-1-11] [Medline: 23961342]

52. Dagöö J, Asplund RP, Bsenko HA, Hjerling S, Holmberg A, Westh S, et al. Cognitive behavior therapy versus interpersonal psychotherapy for social anxiety disorder delivered via smartphone and computer: a randomized controlled trial. J Anxiety Disord 2014 May;28(4):410-417. [doi: 10.1016/j.janxdis.2014.02.003] [Medline: 24731441]

53. Enock PM, Hofmann SG, McNally RJ. Attention Bias Modification Training Via Smartphone to Reduce Social Anxiety: A Randomized, Controlled Multi-Session Experiment. Cogn Ther Res 2014 Mar 4;38(2):200-216. [doi: 10.1007/s10608-014-9606-z]

54. Newman MG, Kenardy J, Herman S, Taylor CB. Comparison of palmtop-computer-assisted brief cognitive-behavioral treatment to cognitive-behavioral treatment for panic disorder. J Consult Clin Psychol 1997 Feb;65(1):178-183. [Medline: 9103747]

55. Agyapong VI, Ahern S, McLoughlin DM, Farren CK. Supportive text messaging for depression and comorbid alcohol use disorder: single-blind randomised trial. J Affect Disord 2012 Dec 10;141(2-3):168-176. [doi: 10.1016/j.jad.2012.02.040] [Medline: 22464008]

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

57. Ly KH, Trüschel A, Jarl L, Magnusson S, Windahl T, Johansson R, et al. Behavioural activation versus mindfulness-based guided self-help treatment administered through a smartphone application: a randomised controlled trial. BMJ Open 2014;4(1):e003440 [FREE Full text] [doi: 10.1136/bmjopen-2013-003440] [Medline: 24413342]

58. Watts S, Mackenzie A, Thomas C, Griskaitis A, Mewton L, Williams A, et al. CBT for depression: a pilot RCT comparing mobile phone vs. computer. BMC Psychiatry 2013;13:49 [FREE Full text] [doi: 10.1186/1471-244X-13-49] [Medline: 23391304]

59. Lappalainen P, Kaipainen K, Lappalainen R, Hoffrén H, Myllymäki T, Kinnunen M, et al. Feasibility of a personal health technology-based psychological intervention for men with stress and mood problems: randomized controlled pilot trial. JMIR Res Protoc 2013;2(1):e1 [FREE Full text] [doi: 10.2196/resprot.2389] [Medline: 23611946]

60. Harrison V, Proudfoot J, Wee PP, Parker G, Pavlovic DH, Manicavasagar V. Mobile mental health: review of the emerging field and proof of concept study. J Ment Health 2011 Dec;20(6):509-524. [doi: 10.3109/09638237.2011.608746] [Medline: 21988230]

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

62. Depp CA, Ceglowski J, Wang VC, Yaghouti F, Mausbach BT, Thompson WK, et al. Augmenting psychoeducation with a mobile intervention for bipolar disorder: a randomized controlled trial. J Affect Disord 2015 Mar 15;174:23-30. [doi: 10.1016/j.jad.2014.10.053] [Medline: 25479050]

63. Wenze SJ, Armey MF, Miller IW. Feasibility and Acceptability of a Mobile Intervention to Improve Treatment Adherence in Bipolar Disorder: A Pilot Study. Behav Modif 2014 Jan 8;38(4):497-515. [doi: 10.1177/0145445513518421] [Medline: 24402464]

64. Ben-Zeev D, Brenner CJ, Begale M, Duffecy J, Mohr DC, Mueser KT. Feasibility, acceptability, and preliminary efficacy of a smartphone intervention for schizophrenia. Schizophr Bull 2014 Nov;40(6):1244-1253. [doi: 10.1093/schbul/sbu033] [Medline: 24609454]

65. Rizvi SL, Dimeff LA, Skutch J, Carroll D, Linehan MM. A pilot study of the DBT coach: an interactive mobile phone application for individuals with borderline personality disorder and substance use disorder. Behav Ther 2011 Dec;42(4):589-600. [doi: 10.1016/j.beth.2011.01.003] [Medline: 22035988]

66. Shapiro JR, Bauer S, Andrews E, Pisetsky E, Bulik-Sullivan B, Hamer RM, et al. Mobile therapy: Use of text-messaging in the treatment of bulimia nervosa. Int J Eat Disord 2010 Sep;43(6):513-519. [doi: 10.1002/eat.20744] [Medline: 19718672]

67. Charness G, Gneezy U, Kuhn MA. Experimental methods: Between-subject and within-subject design. Journal of Economic Behavior & Organization 2012 Jan;81(1):1-8. [doi: 10.1016/j.jebo.2011.08.009]

68. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]

69. Johansson R, Andersson G. Internet-based psychological treatments for depression. Expert Rev Neurother 2012 Jul;12(7):861-9; quiz 870. [doi: 10.1586/ern.12.63] [Medline: 22853793]

70. Frances A. The past, present and future of psychiatric diagnosis. World Psychiatry 2013 Jun;12(2):111-112 [FREE Full text] [doi: 10.1002/wps.20027] [Medline: 23737411]

71. Davison GC. Stepped care: doing more with less? J Consult Clin Psychol 2000 Aug;68(4):580-585. [Medline: 10965633]

72. Hunter J, Schmidt F. Methods of meta-analysis: correcting error and bias in research findings. Thousand Oaks, California: Sage; 2004.

73. Kazdin AE. Mediators and mechanisms of change in psychotherapy research. Annu Rev Clin Psychol 2007;3:1-27. [doi: 10.1146/annurev.clinpsy.3.022806.091432] [Medline: 17716046]

74. Lally P, van Jaarsveld CHM, Potts HWW, Wardle J. How are habits formed: Modelling habit formation in the real world. Eur. J. Soc. Psychol 2010 Oct 16;40(6):998-1009. [doi: 10.1002/ejsp.674]

75. Alkhaldi G, Hamilton FL, Lau R, Webster R, Michie S, Murray E. The Effectiveness of Prompts to Promote Engagement With Digital Interventions: A Systematic Review. J Med Internet Res 2016;18(1):e6 [FREE Full text] [doi: 10.2196/jmir.4790] [Medline: 26747176]

76. Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Med 2010;8:18 [FREE Full text] [doi: 10.1186/1741-7015-8-18] [Medline: 20334633]

77. Moher D, Dulberg CS, Wells GA. Statistical power, sample size, and their reporting in randomized controlled trials. JAMA 1994 Jul 13;272(2):122-124. [Medline: 8015121]

Abbreviations

CBT: cognitive behavioral therapy
EMI: ecological momentary intervention
MHP: mental health professional
N/A: not applicable
PRISMA: Preferred Reporting Items Systematic reviews and Meta-Analyses
RCT: randomized controlled trial

XSL•FO

RenderX

Original Paper

Does a Mobile Phone Depression-Screening App Motivate Mobile Phone Users With High Depressive Symptoms to Seek a Health Care Professional's Help?

Nasser F BinDhim[1], PhD; Eman M Alanazi[1], MHI; Hisham Aljadhey[2,3], PhD; Mada H Basyouni[4], Mphil; Stefan R Kowalski[5,6], PhD; Lisa G Pont[7], PhD; Ahmed M Shaman[2], MClinPharm; Lyndal Trevena[8], PhD; Tariq M Alhawassi[2,3], PhD

[1]College of Health Sciences, Health Informatics, Saudi Electronic University, Riyadh, Saudi Arabia
[2]College of Pharmacy, King Saud University, Riyadh, Saudi Arabia
[3]Medication Safety Research Chair, King Saud University, Riyadh, Saudi Arabia
[4]The Smart Health Project, Riyadh, Saudi Arabia
[5]School of Pharmacy and Medical Sciences, University of South Australia, Adelaide, Australia
[6]Sansom Institute for Health Research, University of South Australia, Adelaide, Australia
[7]Centre for Health Systems and Safety Research, Australian Institute of Health Innovation, Macquarie University, Sydney, Australia
[8]Public Health School, University of Sydney, Sydney, Australia

Corresponding Author:
Nasser F BinDhim, PhD
College of Health Sciences
Health Informatics
Saudi Electronic University
Abu Baker St
Riyadh, 11673
Saudi Arabia
Phone: 966 112613500 ext 1065
Fax: 966 112613578
Email: nbin6641@uni.sydney.edu.au

Abstract

Background: The objective of disease screening is to encourage high-risk subjects to seek health care diagnosis and treatment. Mobile phone apps can effectively screen mental health conditions, including depression. However, it is not known how effective such screening methods are in motivating users to discuss the obtained results of such apps with health care professionals. Does a mobile phone depression-screening app motivate users with high depressive symptoms to seek health care professional advice? This study aimed to address this question.

Method: This was a single-cohort, prospective, observational study of a free mobile phone depression app developed in English and released on Apple's App Store. Apple App Store users (aged 18 or above) in 5 countries, that is, Australia, Canada, New Zealand (NZ), the United Kingdom (UK), and the United States (US), were recruited directly via the app's download page. The participants then completed the Patient Health Questionnaire (PHQ-9), and their depression screening score was displayed to them. If their score was 11 or above and they had never been diagnosed with depression before, they were advised to take their results to their health care professional. They were to follow up after 1 month.

Results: A group of 2538 participants from the 5 countries completed PHQ-9 depression screening with the app. Of them, 322 participants were found to have high depressive symptoms and had never been diagnosed with depression, and received advice to discuss their results with health care professionals. About 74% of those completed the follow-up; approximately 38% of these self-reported consulting their health care professionals about their depression score. Only positive attitude toward depression as a real disease was associated with increased follow-up response rate (odds ratio (OR) 3.2, CI 1.38-8.29).

Conclusions: A mobile phone depression-screening app motivated some users to seek a depression diagnosis. However, further study should investigate how other app users use the screening results provided by such apps.

XSL·FO
RenderX

Introduction

Great computational and storage abilities as well as proximity to users potentially make mobile phone apps excellent health research tools that are capable of delivering complex health interventions [1]. In addition, research participants can be recruited directly from app stores [2,3]. Hence, various studies have explored the use of mobile phone apps locally and cross-country for health research, including cross-sectional studies [2,3], observational studies [4], and randomized controlled trials [5,6].

Mobile phone and Mental Health Screening

Current research on apps for screening and monitoring mental health has shown feasibility across diverse ranges of mental health conditions, including depression [3], bipolar disorder [7,8], anxiety disorders [9,10], and substance abuse disorders [2,5,6-11]. Furthermore, recent studies suggest that mobile phone ownership is very common among mental health patients and they have a strong interest in using mobile phones to monitor their mental health [12,13]. However, there is still very limited evidence regarding the efficacy of mobile phone-delivered mental health interventions or screening tools [14]. The limited quantity of studies conducted in this domain might be due to the lack of feasibility and low confidence that these interventions and/or screening tools will reach the targeted populations [3], and may be related to the fact that mobile phone technology is still nascent compared with other delivery channels for health interventions.

In a recent cross-sectional study, 8241 users from 66 countries from Apple's App Store downloaded a depression-screening app [3]. A high percentage (73.9%) of app downloaders also submitted responses to the screening questionnaire [3], with 25.7% reporting that they had previously been diagnosed with depression [3]. Using two cutoff thresholds of the Patient Health Questionnaire (PHQ-9) depression screening tool, it was found that a large number of participants had high depressive symptoms yet were undiagnosed [3-15]. The studied app also reached various groups of people who were not previously diagnosed with depression yet had high depressive symptoms, and the app was able to reach a group of participants at risk of suicide [3]. In another study in South Korea, 27,159 participants were screened for bipolar spectrum disorders within a few months [8].

Mobile Phone Apps' Potential in Mental Health Screening and Monitoring

Active data, in the form of questionnaires for mental health screening or monitoring, can bring clinical assessments from outside of the health care setting into the real-life environment, lived and experienced by patients [16]. In addition, there are dozens of mobile phone apps that provide mental health and depression screening, but few advise users to discuss results with their health care professionals and most do not provide any guidance about how to use such results. Due to the wide popularity of screening tests on mobile phone apps and on the Internet, people are likely to access and use apps, especially if aware of their depressive symptoms. However, once they are aware of the potential seriousness of symptoms, it is critical they take appropriate action and seek professional follow-up by discussing the results with a health care professional. However, it is not yet known how effective such screening methods are in motivating users to seek appropriate professional help. As the objective of screening for disease is to discover the undiagnosed problem so that the users can be placed under treatment [17], establishing an association between mobile phone self-screening for mental health conditions and the actions taken by the users based on their screening results is a critical step in assessing the feasibility, efficacy, and cost-effectiveness of such screening and monitoring methods.

The aim of this study was to address the following question: Does a mobile phone depression-screening app motivate users with high depressive symptoms to seek health care professional advice?

Methods

Design

This study was a single-cohort, prospective, observational study of a free mobile phone depression app that was developed in English utilizing the "Health Monitor" app template [18] and was released on Apple's App Store. The users of the Apple's App Store from any country can download the app after consenting to provided participant information; however, we limited the app availability to 5 countries: Australia, Canada, New Zealand (NZ), the United Kingdom (UK), and the United States (US). We selected these countries based on their high download and response rates in a previous feasibility study [3]. Users from other countries were excluded because, as provided by a previous study, the app stores function of limiting app users to specific countries in not fully accurate [2]. The research ethics committee at King Saud University approved this study.

Participants

Apple App Store users aged 18 or above, from the 5 nominated countries were recruited directly via the app's download page in the Apple App Store. The consent and participants' information are summarized in the app download page (Figure 1) and are included in the "about" section of the app. The studied app was published during the recruitment period (January 25 to March 25, 2015) to the Apple App Store, which was the main portal for advertising this study's app. In addition, to boost the recruitment process, we advertised the app with demographic targeting (by country) using an in-app advertisement, that results in the app ad being displayed to Apple iPhone users while they are using other apps.

XSL•FO

RenderX

Figure 1. Screenshot of the app download page.
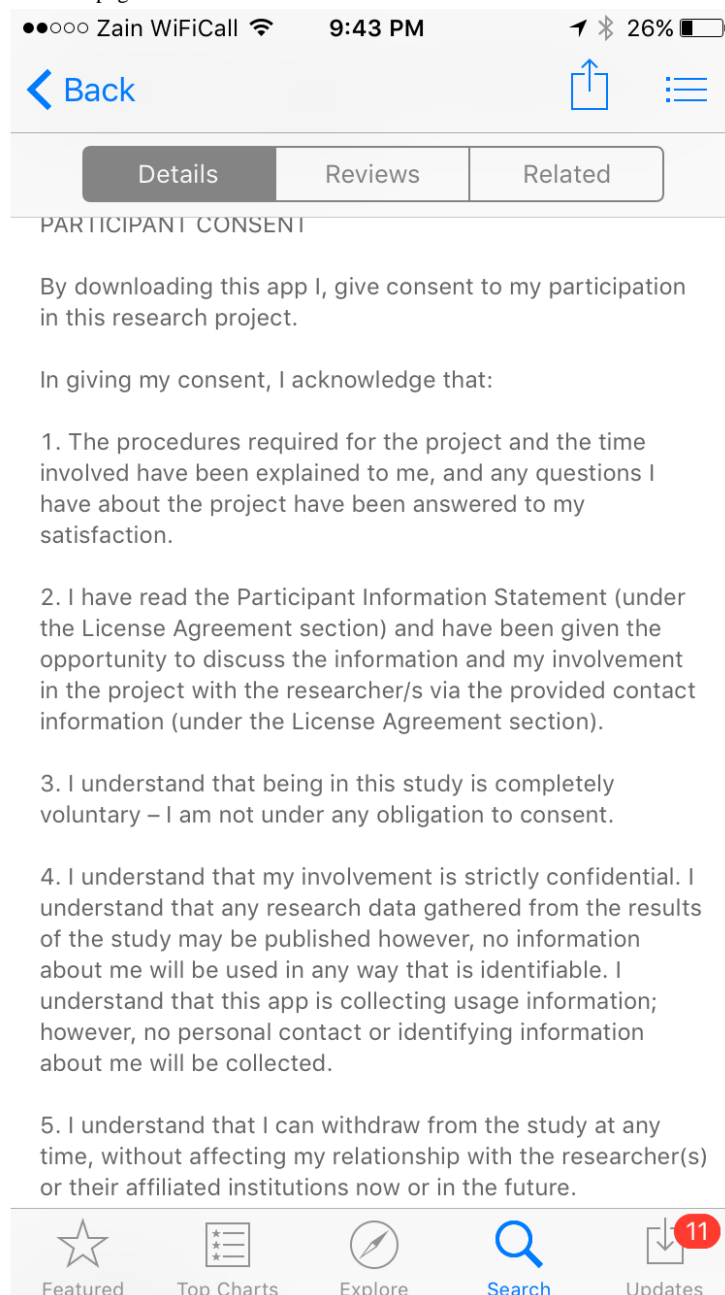
Recruitment and Data Collection

When a participant downloaded the app and ran it for the first time, the app assigned a unique device identifier to the user's device and registered it in our secure online research database. The identifier will not change even if the user deletes the app or resets the device. This allowed anonymous data collection and prevented duplicate enrollments. After submitting demographic and baseline data on the first screen of the app, which included educational level, employment and income status, and other health-related characteristics such as chronic conditions, (Figure 2) the app users were able to complete the PHQ-9 on the second screen, including previous depression diagnosis and treatment as well as other depression risk factors, and were able to see their depression screening score. The PHQ-9 was selected over other depression screening tools because (1) it has been validated for use among various age groups [19-21] and (2) there appears to be strong correlation between scores reported from the app and scores reported on paper, with app-collected scores 3.02 (SD 2.25) points higher on average [22].

If the users get a score of 11 or more, the app recommends they discuss the test score with a health care professional. A threshold score of 11 or above was selected based on the literature, which, in pooled estimates of 10 studies had the best trade-off between sensitivity, 0.89 (95% CI 0.75 to 0.96), and specificity, 0.89 (95% CI 0.79 to 0.94) [23]. It is important to note that our app did not provide a diagnosis to the user as our main intention was to replicate the available mental health screening tests on the mobile phone app stores or the Internet. The classification of participants based on the PHQ-9 cutoff 11 was used internally and was not communicated to the participants.

In addition, all of the participants with PHQ-9 scores of 11 or above who did not indicate a previous diagnosis of depression were listed for automated follow-up after 1 month using push notifications. Push notification permits the database server to send messages to a specific user (similar to mobile short message service) at specified times or after a specific task, free of cost [24]. The 1-month follow-up period was selected arbitrarily because there are no specific guidelines regarding a definitive follow-up period but it struck a balance between clinical responsibility and effective push notification practicality. However, at the follow-up, we asked participants if they had discussed the results the app provided with a health care professional, and if so, that professional had diagnosed them with depression or not. We sent 5 push notification reminders over a 10-day period to help increase the follow-up response rate.

Figure 2. Screenshots of the app.

Data Analysis

The mean and SD of the quantitative variables were presented if they had a normal distribution, or the median and range were presented, as appropriate, and compared using *t* tests. We presented the categorical variables as percentages and CIs, and compared using Pearson's chi-square test. As this study used automated electronic data collection, there were no missing values in the baseline data; the app also includes a data integrity check to prevent users from entering invalid data (such as a maximum age of 99). We used logistic regression to explore factors associated with response rate to the follow-up and with seeking professional help.

Results

Uptake and Demographics

In 2 months, 3984 users downloaded the app and 2689 out of 3984 (67.49%) completed the PHQ-9 screening. Of those who completed the screening, 151 users were not eligible for follow-up because they were not from the 5 countries targeted in this study and we were unable to identify from which country were. Thus, 2538 participants were included in this study and were eligible for follow-up. Table 1 shows the participants' demographics.

Table 1. Demographics.

Characteristics		Total n (%)
Age (median range) (years)		27.0 (18-75)
Sex		
	Female	1143 (45.0)
	Male	1395 (54.9)
Country		
	Australia	283 (11.2)
	Canada	541 (21.3)
	United Kingdom	1108 (43.7)
	United States	497 (19.6)
	New Zealand	109 (4.3)
Education		
	Above high school	1177 (46.4)
	High school or less	1361 (53.6)
Income level		
	Less than US $20,000/year	1276 (50.3)
	US $21,000-49,000/year	720 (28.4)
	More than US $50,000/year	542 (21.4)
Employment status		
	Employed	1151 (45.4)
	Self-employed	285 (11.2)
	Student	593 (23.4)
	Unemployed	509 (20.0)

Table 2. Country-based prevalence of undiagnosed higher risk of depression using the PHQ-9 threshold of 11.

	Country					
	Australia	Canada	New Zealand	United Kingdom	United States	Total
Low depressive symptoms (PHQ 11)	206 (67.3%)	390 (71.8%)	66 (72.5%)	774 (69.9%)	368 (74.0%)	1804 (70.9%)
High depressive symptoms (PHQ 11)	100 (32.7%)	153 (28.2%)	25 (27.5%)	334 (30.1%)	129 (26.0%)	741 (29.1%)

Depression Screening and Follow-Up

There were 741 participants out of 2538 (29.1%) with high depressive symptoms (PHQ ≥ 11) as shown in Table 2, of which 419 (56.5%) had been previously diagnosed with depression.

A total of 322 participants, therefore, fulfilled the study criteria for follow-up reminder messages. Of those who followed up, 239 out of 322 (74.2%) completed the follow-up questions. Addressing the principal aim of this study, 91 out of 239 (38%) self-reported consulting their health care professionals about the depression score provided via the app.

Broken down by country of origin, 26 out of 53 (49%) Australian participants, who followed up, reported consulting their health care professionals regarding their PHQ-9 scores, compared with 23 out of 50 (46%) from Canada, 28 out of 85 (33%) from the United Kingdom, and only 14 out of 51 (27%) from the United States, with no differences between the countries (χ^2_3= 7.4, P=.059). Eventually, 27 out of 91 (29%) participants self-reported being diagnosed with depression as a result of the consultation with health care professionals.

Logistic regression analysis incorporating demographics (Table 1) and health characteristics variables (Table 3) identified only positive attitude toward depression as a real disease as being associated with increased follow-up response rate OR 3.2, CI 1.38–8.29.

XSL•FO
RenderX

Table 3. Health characteristics.

Characteristics		Total
		n (%)
Alcohol consumption	Never	1218 (48.0)
	1-2 standard drinks/occasion	648 (25.6)
	3-4 standard drinks/occasion	299 (11.8)
	5+ standard drinks/occasion	373 (14.6)
Chronic disease		
	No	1747 (68.8)
	Yes	791 (31.2)
Depression diagnosis		
	No	1934 (76.2)
	Yes	604 (23.8)
High depressive symptoms		
	PHQ-9 cutoff of 11 or above	741 (29.2)
	PHQ-9 less than cutoff of 11	1797 (70.8)
Cigarette smoking		
	No	1725 (67.9)
	Yes, 10 cigarettes or less/day	352 (13.9)
	Yes, 11-20 cigarettes/day	310 (12.2)
	Yes, 21 cigarettes or more/day	151 (5.9)
Attitude toward depression (Depression is a real illness)		
	Strongly agree/Agree	2184 (86.1)
	Neutral	262 (10.3)
	Strongly disagree/Disagree	92 (3.6)
Attitude toward anti-depressant medications (Do they help restore normal level of functioning?)		
	Strongly agree/Agree	1107 (43.6)
	Neutral	1089 (42.9)
	Strongly disagree/Disagree	342 (13.4)
Attitude toward counseling (Help restore normal level of functioning)		
	Strongly agree/Agree	1407 (55.4)
	Neutral	843 (33.2)
	Strongly disagree/Disagree	288 (11.4)

Discussion

In this study, 2538 participants from Australia, Canada, New Zealand, the United Kingdom, and the United States, completed the PHQ-9 depression screening using a mobile phone app. Of the respondents, 322 participants with high depressive symptoms who had not previously been diagnosed with depression were directed via the app to seek health care professional advice. The app also sent a follow-up message after 1 month using a mobile phone push notification asking users if they had sought health professional advice for the depression score they received from the app. Approximately 74% (239 out of 322) of users who

scored highly on the app completed the follow-up, of which 38% (91 out of 239) had self-reported that they had consulted their health care professionals about the depression score provided via the app. The highest proportions of participants who had consulted a health care professional were from Australia and Canada.

Depression Screening via Mobile Phone App and Motivation to Seek Help

This study demonstrated that mobile phone users from a variety of countries were willing to use the depression-screening app and some acted on the results. More than one-third of the

XSL•FO
RenderX

follow-up respondents acted on the recommendation provided via the app and 29% (27 out of 91) self-reported that they had received a diagnosis of depression from their health care professional at follow-up. This shows that mobile phone depression screening can influence users to discuss results obtained from a mobile phone screening test. However, we still do not know how the majority who did not sought help benefited from or used screening results.

The findings of this study on the ability of depression screening to motivate some of those screened to discuss the results with health care professionals are well matched with previous studies [25]. Although in this study we relied on the user to initiate the processes of seeking the health care professional help to interpret the results, such mobile phone mental health screening might be more effective if linked to electronic health records so that clinicians can view it to enhance the communication process. This should also follow the US Preventive Services Task Force's recent recommendations about screening for depression in the general adult population, which also recommends adequate systems in place to ensure accurate diagnosis, effective treatment, and appropriate follow-up [26].

Depression Prevalence and Usage of the Depression-Screening App

Different studies have shown a strong relationship between socioeconomic status and the prevalence of depression. In this study, the majority of the participants who used the depression-screening app were in a lower-income demographic and earned less than US $20,000/year [27]. Although the prevalence of depression in the United States is supposedly high and has been reported to affect approximately 1 in 5 people [27,28], there were fewer participants from the United States, 497 (19.6%), than from the United Kingdom, 1108 (43.66%). There was a similar finding, 109 (4.3%) participants in the New Zealand subgroup that had previously been reported to have a high prevalence of depression had a lower response rate in this study. Thus, variations in mobile phone use and app ranking in each country might be the major causes of such variation [29].

Variation in Seeking Health Care Professional Advice

Australia and Canada had the highest proportions of users 49% (26 out of 53) and 46% (23 out of 50) respectively, who scored high on the app and reported going on to discuss the results with health care professionals. One of the possible reasons why more participants from Australia and Canada sought professional help could have been access to low-cost subsidized government-run health care systems. As previously mentioned, the majority of the participants had incomes of less than $20K/year, 1276 out of 2538 (50.28%). In the United States, the average direct annual per patient costs were $10,402 for bipolar patients and $7494 for depressed patients [30]. Therefore, providing free or affordable health care costs might be a major factor that motivates people with depression to seek professional help after general population screening. In addition, employees in the United States who are treated for depression incur annual per capita health and disability costs of $5415, which is significantly more than other diseases like hypertension, and treatment associated with a mean of 9.86 annual sick days, which is significantly more than other conditions [31].

Implications

There are various mobile phone apps in app stores that provide mental health or other health screening assessment tools (such as cancer screeners). However, the impact of such apps on consumer decisions is unknown. Moreover, there is no universal app store policy to responsibly direct consumers' behavior regarding specific therapy [32]. It is recommended that app stores implement a responsible policy to force medical apps to declare that screening results obtained should be discussed with appropriate health care professionals. This criterion should also be included in the quality evaluation of any mobile phone health app. However, while positive screening results via mobile phone apps may stress a user, false negatives may also harm users, demotivating them from seeking health care advice.

Mental health screening apps need to implement functions to better motivate users to act on the provided screening results and encourage them to discuss the results with health care professionals. This study demonstrated that 148 out 239 (62%) respondents who demonstrated significant depressive symptoms did not discuss their results with a health care professional. The screening function alone might therefore not be sufficiently effective in motivating users to seek professional assistance. Improving this utility is critical when designing and implanting mobile phone technology for assistance with health care provision.

In a previous mobile phone depression-screening study, (2642/6089) 43.38% of participants completed the PHQ-9 questionnaire an average 5.3 times in the 4-month study period, with a depression score above 11 at the first test being associated with multiple PHQ-9 completions [3]. In another case study, 13 patients with major depressive disorder used a simple mobile phone app to answer 3 randomly sampled questions from the PHQ-9 survey 3 times per day for the duration of 1 month [33]. Given the feasibility of using mobile phone apps to monitor depressive symptoms, comparing a patient's current responses to previous responses might help mental health practitioners make informed decisions [34]. Finally, although 2184 out of 2538 (86.05%) participants recorded a positive attitude toward depression as a real disease, only 1107 out of 2538 (43.61 %) had a positive attitude toward antidepressant medications' ability to restore normal level of functioning. Future mental health mobile phone interventions may need to consider adding a component to improve attitude toward antidepressant medications.

Limitations

This study focused on simulating the real-world use of such screening tools. It deliberately recruited participants the typical way mobile phone users will seek apps and provided new evidence that relevant users in various countries seek and use mobile phone-based mental health interventions. This process limited the ability to validate the self-reported data, which is less rigorous than clinically validated data. Another limitation of this study is that the passive recruitment strategy might have drawn in participants more motivated to complete the screening and the follow-up, perhaps those aware of their mental health problem. To test the feasibility of using push notifications to follow up with participants, we used a short set of follow-up

questions. Therefore, this study did not provide details about the type of depression diagnosis or who provided the depression diagnosis (eg, a general practitioner or specialist). It is also unlikely that the population included in this study is representative of the depression prevalence in the population, as users had to have a mobile phone and be active app users. Moreover, the cohort design lacked a comparison or control group, which can provide results that are more rigorous. Finally, the demographics of users in this study do not reflect the national averages. For example, in the United States, depression rates were higher in 40-59 year olds and women. [35] However, in this study most of the participants were younger and more participants were male than females.

Conclusion

Previous studies have confirmed the feasibility of depression screening using mobile phone apps in various countries; however, it was unknown if such screening could motivate users to discuss the obtained results with health care professionals, and lead to clinical diagnosis. This study showed that a mobile phone depression-screening app could motivate some users to discuss the obtained results of such tests with health care professionals for further diagnosis and management. However, further study should investigate how other app users use the screening results provided by depression-screening apps.

Authors' Contributions

NFB designed the study and analyzed the data. MHB managed the data collection and the app design. EMA and NFB wrote the first draft of the manuscript. All authors contributed to and have approved the final manuscript.

Conflicts of Interest

None declared.

References

1. BinDhim NF, Trevena L. There's an App for That: A Guide for Healthcare Practitioners and Researchers on Smartphone Technology. Online J Public Health Inform 2015 May;7(2):e218 [FREE Full text] [doi: 10.5210/ojphi.v7i2.5522] [Medline: 26392848]

2. BinDhim NF, McGeechan K, Trevena L. Who Uses Smoking Cessation Apps? A Feasibility Study Across Three Countries via Smartphones. JMIR Mhealth Uhealth 2014;2(1):e4 [FREE Full text] [doi: 10.2196/mhealth.2841] [Medline: 25098439]

3. BinDhim NF, Shaman AM, Trevena L, Basyouni MH, Pont LG, Alhawassi TM. Depression screening via a smartphone app: cross-country user characteristics and feasibility. J Am Med Inform Assoc 2015 Jan;22(1):29-34 [FREE Full text] [doi: 10.1136/amiajnl-2014-002840] [Medline: 25326599]

4. Patel V, Nowostawski M, Thomson G, Wilson N, Medlin H. Developing a smartphone 'app' for public health research: the example of measuring observed smoking in vehicles. J Epidemiol Community Health 2013 May;67(5):446-452. [doi: 10.1136/jech-2012-201774] [Medline: 23443959]

5. Gustafson DH, McTavish FM, Chih M, Atwood AK, Johnson RA, Boyle MG, et al. A smartphone application to support recovery from alcoholism: a randomized clinical trial. JAMA Psychiatry 2014 May;71(5):566-572 [FREE Full text] [doi: 10.1001/jamapsychiatry.2013.4642] [Medline: 24671165]

6. Bricker JB, Mull KE, Kientz JA, Vilardaga R, Mercer LD, Akioka KJ, et al. Randomized, controlled pilot trial of a smartphone app for smoking cessation using acceptance and commitment therapy. Drug Alcohol Depend 2014 Oct 1;143:87-94. [doi: 10.1016/j.drugalcdep.2014.07.006] [Medline: 25085225]

7. Faurholt-Jepsen M, Frost M, Vinberg M, Christensen EM, Bardram JE, Kessing LV. Smartphone data as objective measures of bipolar disorder symptoms. Psychiatry Res 2014 Jun 30;217(1-2):124-127. [doi: 10.1016/j.psychres.2014.03.009] [Medline: 24679993]

8. Woo YS, Bahk W, Hong J, Yoon B, Hwang T, Kim M, et al. Use of a smartphone application to screen for bipolar spectrum disorder in a community sample. Health Informatics J 2015 Jul 17:NA Epub ahead of print(forthcoming). [doi: 10.1177/1460458215589601] [Medline: 26187988]

9. Pramana G, Parmanto B, Kendall PC, Silk JS. The SmartCAT: an m-health platform for ecological momentary intervention in child anxiety treatment. Telemed J E Health 2014 May;20(5):419-427 [FREE Full text] [doi: 10.1089/tmj.2013.0214] [Medline: 24579913]

10. BinDhim NF, Shaman A, Alhawassi T. Confirming the one-item question Likert Scale to measure anxiety. Internet J Epidemiol 2013;11:1-2.

11. BinDhim N. Sydney escholarship repository. Sydney: University of Sydney; 2015. Smartphone applications for health consumers: the example of smoking URL: http://ses.library.usyd.edu.au:80/handle/2123/12708 [accessed 2016-06-03] [WebCite Cache ID 6hzH0Y3Dj]

XSL·FO
RenderX

12. Torous J, Chan SR, Yee-Marie TS, Behrens J, Mathew I, Conrad EJ, et al. Patient Smartphone Ownership and Interest in Mobile Apps to Monitor Symptoms of Mental Health Conditions: A Survey in Four Geographically Distinct Psychiatric Clinics. JMIR Ment Health 2014;1(1):e5 [FREE Full text] [doi: 10.2196/mental.4004] [Medline: 26543905]

13. Torous J, Friedman R, Keshavan M. Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. JMIR Mhealth Uhealth 2014;2(1):e2 [FREE Full text] [doi: 10.2196/mhealth.2994] [Medline: 25098314]

14. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. J Med Internet Res 2013;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]

15. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x]

16. Torous J, Staples P, Onnela J. Realizing the potential of mobile mental health: new methods for new data in psychiatry. Curr Psychiatry Rep 2015 Aug;17(8):602. [doi: 10.1007/s11920-015-0602-0] [Medline: 26073363]

17. Jungner G, Wilson JMG. Principles and practice of screening for disease. In: Public Health Papers; no. 34. Geneva: WHO; 1968.

18. BinDhim N. Smart Health project. 2012. Health monitor project URL: http://shproject.net/index. php?option=com_content&view=article&id=3 [WebCite Cache ID 6fnjMgMlx]

19. Richardson LP, McCauley E, Grossman DC, McCarty CA, Richards J, Russo JE, et al. Evaluation of the Patient Health Questionnaire-9 Item for detecting major depression among adolescents. Pediatrics 2010 Dec;126(6):1117-1123 [FREE Full text] [doi: 10.1542/peds.2010-0852] [Medline: 21041282]

20. Allgaier A, Pietsch K, Frühe B, Sigl-Glöckner J, Schulte-Körne G. Screening for depression in adolescents: validity of the patient health questionnaire in pediatric care. Depress Anxiety 2012 Oct;29(10):906-913. [doi: 10.1002/da.21971] [Medline: 22753313]

21. Phelan E, Williams B, Meeker K, Bonn K, Frederick J, Logerfo J, et al. A study of the diagnostic accuracy of the PHQ-9 in primary care elderly. BMC Fam Pract 2010;11:63 [FREE Full text] [doi: 10.1186/1471-2296-11-63] [Medline: 20807445]

22. Torous J, Staples P, Shanahan M, Lin C, Peck P, Keshavan M, et al. Utilizing a Personal Smartphone Custom App to Assess the Patient Health Questionnaire-9 (PHQ-9) Depressive Symptoms in Patients With Major Depressive Disorder. JMIR Ment Health 2015 Mar;2(1):e8 [FREE Full text] [doi: 10.2196/mental.3889] [Medline: 26543914]

23. Manea L, Gilbody S, McMillan D. Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. CMAJ 2012 Feb 21;184(3):E191-E196 [FREE Full text] [doi: 10.1503/cmaj.110829] [Medline: 22184363]

24. BinDhim NF, McGeechan K, Trevena L. Assessing the effect of an interactive decision-aid smartphone smoking cessation application (app) on quit rates: a double-blind automated randomised control trial protocol. BMJ Open 2014;4(7):e005371 [FREE Full text] [doi: 10.1136/bmjopen-2014-005371] [Medline: 25037644]

25. Frye MA, Calabrese JR, Reed ML, Wagner KD, Lewis L, McNulty J, Hirschfeld Robert M A. Use of health care services among persons who screen positive for bipolar disorder. Psychiatr Serv 2005 Dec;56(12):1529-1533. [doi: 10.1176/appi.ps.56.12.1529] [Medline: 16339614]

26. Siu AL, US Preventive Services Task Force (USPSTF), Bibbins-Domingo K, Grossman DC, Baumann LC, Davidson KW, et al. Screening for Depression in Adults: US Preventive Services Task Force Recommendation Statement. JAMA 2016 Jan 26;315(4):380-387. [doi: 10.1001/jama.2015.18392] [Medline: 26813211]

27. Lorant V, Croux C, Weich S, Deliège D, Mackenbach J, Ansseau M. Depression and socio-economic risk factors: 7-year longitudinal population study. Br J Psychiatry 2007 Apr;190:293-298 [FREE Full text] [doi: 10.1192/bjp.bp.105.020040] [Medline: 17401034]

28. Kessler RC, Bromet EJ. The epidemiology of depression across cultures. Annu Rev Public Health 2013;34:119-138 [FREE Full text] [doi: 10.1146/annurev-publhealth-031912-114409] [Medline: 23514317]

29. BinDhim NF, Freeman B, Trevena L. Pro-smoking apps: where, how and who are most at risk. Tob Control 2015 Mar;24(2):159-161. [doi: 10.1136/tobaccocontrol-2013-051189] [Medline: 24046212]

30. Stensland MD, Jacobson JG, Nyhuis A. Service utilization and associated direct costs for bipolar disorder in 2004: an analysis in managed care. J Affect Disord 2007 Aug;101(1-3):187-193. [doi: 10.1016/j.jad.2006.11.019] [Medline: 17254637]

31. Druss BG, Rosenheck RA, Sledge WH. Health and disability costs of depressive illness in a major U.S. corporation. Am J Psychiatry 2000 Aug;157(8):1274-1278. [doi: 10.1176/appi.ajp.157.8.1274] [Medline: 10910790]

32. BinDhim NF, Trevena L. Health-related smartphone apps: regulations, safety, privacy and quality: Figure 1. BMJ Innov 2015 Mar 05;1(2):43-45. [doi: 10.1136/bmjinnov-2014-000019]

33. Torous J, Staples P, Shanahan M, Lin C, Peck P, Keshavan M, et al. Utilizing a Personal Smartphone Custom App to Assess the Patient Health Questionnaire-9 (PHQ-9) Depressive Symptoms in Patients With Major Depressive Disorder. JMIR Ment Health 2015 Mar;2(1):e8 [FREE Full text] [doi: 10.2196/mental.3889] [Medline: 26543914]

34. Torous J, Staples P, Onnela J. Realizing the potential of mobile mental health: new methods for new data in psychiatry. Curr Psychiatry Rep 2015 Aug;17(8):602. [doi: 10.1007/s11920-015-0602-0] [Medline: 26073363]

XSL·FO
RenderX

35.   Pratt LA, Brody DJ. National Center for Health Statistics Data Brief. Atlanta: Centers for Disease Control and Prevention; 2008. Depression in the United States household population, 2005-2006 URL: http://www.cdc.gov/nchs/data/databriefs/db07.pdf [accessed 2016-06-03] [WebCite Cache ID 6hzKGUMci]

Original Paper

Social Media in Professional Medicine: New Resident Perceptions and Practices

Cedric Lefebvre[1*], MD; Jason Mesner[1*], MD; Jason Stopyra[1*], MD; James O'Neill[1*], MD; Iltifat Husain[1*], MD; Carol Geer[2*], MD; Karen Gerancher[3*], MD; Hal Atkinson[4*], MD; Erin Harper[1*], MSHS; William Huang[5*], MD; David M Cline[1*], MD

[1]Wake Forest School of Medicine, Department of Emergency Medicine, Winston Salem, NC, United States

[2]Wake Forest School of Medicine, Department of Radiology, Winston Salem, NC, United States

[3]Wake Forest School of Medicine, Department of Obstetrics and Gynecology, Winston Salem, NC, United States

[4]Wake Forest School of Medicine, Department of Internal Medicine, Winston Salem, NC, United States

[5]Wake Forest School of Medicine, Department of Dermatology, Winston Salem, NC, United States

[*]all authors contributed equally

Corresponding Author:
Cedric Lefebvre, MD
Wake Forest School of Medicine
Department of Emergency Medicine
1 Medical Center Boulevard
Winston Salem, NC, 27157
United States
Phone: 1 336 716 2190
Fax: 1 336 716 1705
Email: clefebvr@wakehealth.edu

Abstract

Background: For younger generations, unconstrained online social activity is the norm. Little data are available about perceptions among young medical practitioners who enter the professional clinical arena, while the impact of existing social media policy on these perceptions is unclear.

Objective: The objective of this study was to investigate the existing perceptions about social media and professionalism among new physicians entering in professional clinical practice; and to determine the effects of formal social media instruction and policy on young professionals' ability to navigate case-based scenarios about online behavior in the context of professional medicine.

Methods: This was a prospective observational study involving the new resident physicians at a large academic medical center. Medical residents from 9 specialties were invited to participate and answer an anonymous questionnaire about social media in clinical medicine. Data were analyzed using SAS 9.4 (Cary, NC), chi-square or Fisher's exact test was used as appropriate, and the correct responses were compared between different groups using the Kruskal–Wallis analysis of variance.

Results: Familiarity with current institutional policy was associated with an average of 2.2 more correct responses (P=.01). Instruction on social media use during medical school was related to correct responses for 2 additional questions (P=.03). On dividing the groups into no policy exposure, single policy exposure, or both exposures, the mean differences were found to be statistically significant (3.5, 7.5, and 9.4, respectively) (P=.03).

Conclusions: In this study, a number of young physicians demonstrated a casual approach to social media activity in the context of professional medical practice. Several areas of potential educational opportunity and focus were identified: (1) online privacy, (2) maintaining digital professionalism, (3) safeguarding the protected health information of patients, and (4) the impact of existing social media policies. Prior social media instruction and/or familiarity with a social media policy are associated with an improved performance on case-based questions regarding online professionalism. This suggests a correlation between an instruction about online professionalism and more cautious online behavior. Improving the content and delivery of social media policy may assist in preserving institutional priorities, protecting patient information, and safeguarding young professionals from online misadventure.

XSL•FO
RenderX

Introduction

The use of social media by members of professional groups has grown substantially in the past decade. In medical practice, online activities offer potential benefits to providers, patients, and the profession [1]; and present unique challenges for users and their employers. For younger generations, unconstrained online social activity is the norm, but as these tech-savvy users transition into professional life, their attitude toward online security and professionalism may not match that of their employers or their professions. Many health care institutions embrace social media policies, although it is not clear whether these policies address misperceptions among social media users or mitigate potential pitfalls [2]. Despite the existence of social media policies, professionals continue to find themselves in ethical, professional, and/or legal trouble [3-9]. Little data are available about the perceptions and social media practices among the young medical practitioners entering the professional clinical arena. The impact of the existing social media policy on the attitude of these professionals is also unclear.

Methods

This was a prospective observational study involving the new resident physicians in a large academic medical center. Residents were invited to participate and answer an anonymous questionnaire regarding their personal use of social media and their perceptions of appropriateness while using social media in the context of clinical medicine. Medical residents from the departments of internal medicine, family medicine, general surgery, emergency medicine, neurology, anesthesiology, neurosurgery, radiology, and obstetrics and gynecology were included. An approval was obtained from the Institutional Review Board (IRB). The 30-item questionnaire covered the following: self-reported instruction on social media use in medical school (1); self-reported familiarity with current institutional social media policy (1); case-scenario questions regarding specific tenets of the current institutional social media policy (15); understanding of social media account control options (5); prior social media account closures (2); prior unprofessional posts (2); current social media use (2); and demographics (2).

On the basis of the correct responses to the questions regarding specific tenets of the current institutional social media policy, a score was calculated for each participant (0 to a maximum of 15). Questionnaires were delivered electronically to each resident by their respective program directors during an orientation period (3-5 days before their clinical start date). In accordance with the medical center's usual orientation procedures, each resident received printed information about the institution's social media policy from the Graduate Medical Education (GME) office several days before starting clinical duties and receiving the invitation to attempt this questionnaire. The policy was also available for review on the institution's Website. The study period was June 25-July 15, 2015. Upon

completion of the questionnaire, the subjects were invited to review the recommended answers and explanations developed by the expert consensus. These explanations were based on the existing institutional social media policy and social media best practices. Data were analyzed using SAS 9.4 (Cary, NC), categorical variables were compared using chi-square or Fisher's exact test as appropriate, and the correct responses to questions regarding specific tenets of the current institutional social media policy were compared between different groups using the Kruskal–Wallis analysis of variance.

Results

Of 124 invited subjects, 70 participated in the questionnaire (response rate of 56%). Of the 70 questionnaires, 13 contained incomplete answers. Most incomplete answers were observed in the latter half of the questionnaire. The highest rate of incompletion was associated with questions 24-31 (see Multimedia Appendix 1). About 98% of the respondents were born between 1979 and 1998. The majority of respondents (60/70, 86%) then reported having a social media account (eg, Facebook, Twitter) while 51% (36/70) reported using an image messaging app (eg, Snapchat, Flickr).

The cohort of residents familiar with the current institutional policy on social media was 29% (20/70). Familiarity with the current institutional policy on social media had a statistically significant effect on response tendencies for 2 individual questionnaire items. Of those familiar with the policy who answered a question about posting a picture of a colleague, a greater percentage (63%, 12/19) selected the correct answer than those who were unfamiliar with the policy (32%, 14/44, P=.0204). On a question about posting an image of a patient who is "unidentifiable," 100% of residents who reported familiarity with the social media policy answered correctly ("never okay"), compared with only 78% (32/41) of residents who were unfamiliar with the policy (P=.0237).

A majority of the participating residents (67%, 47/70) received formal instruction about social media use previously during medical school. Exposure to social media policies in medical school was associated with statistically significant response trends for 2 questionnaire items. When participants were asked about accepting a social network invitation from a professional acquaintance with whom no significant social or professional relationship exists, 100% of respondents with no medical school instruction about social media answered incorrectly compared with 72% (29/40) of subjects who reported prior instruction during medical school (P=.0125). In response to a question about interacting with patients via social media, 67% (27/40) of interns with prior medical school instruction answered correctly ("never okay"). Among interns with no prior medical school instruction on social media, only 35% (6/17) submitted a correct response regarding communication with patients via social media (P=.0243).

Participants were grouped according to their report of prior medical school instruction on social media and their familiarity

with the current institutional social media policy. The results are summarized in Tables 1 and 2. The potential correct response total was 15; the actual range of correct responses was 0-13. Familiarity with current institutional policy was associated with an average of additional 2.2 more correct responses. Instruction on social media use during medical school similarly was associated with correct test scores for an average of 2 additional questions (see Table 1). On dividing the groups into no positive factor, either (single) positive factor, or both factors, the mean differences were found to be statistically significant (3.5, 7.5, and 9.4, respectively), using the Kruskal–Wallis test (see Table 2).

Of the participants, 39% (27/70) believed that digital information, once posted, could be permanently deleted, while 21% (15/70) were not sure. When the participants were asked,

whether logging into a personal social media account while on duty in a patient care area is acceptable (not permitted by institutional policy), 27% (17/64) selected incorrect answers ("it depends" or "always"), while 5% (3/64) were "unsure." In response to a question about posting pictures of a patient's discrete physical finding on social media with no obvious way to identify the patient (not permitted by hospital policy), 15% (9/60) described this practice as "always ok, it depends, or unsure." Regarding interprofessional interaction on social media, 25% (14/57) of subjects indicated that it is "always okay" to accept a social media invitation (eg, "friend request") from a nurse with whom there is otherwise no social relationship. When the participants were asked whether it is acceptable to interact with patients on social media, which constitutes conduct that is discouraged by the social media policy, 33% (19/57) participants answered "it depends."

Table 1. Average correct score stratified by institutional social media policy familiarity and medical school instruction on social media.

Parameter	N	Mean correct[b]	95% CI	P-value[a]
Familiar with institutional social media policy	20	9.0	7.6-10.4	.01
Not familiar with institutional social media policy	49	6.8	5.8-7.8	
Instruction on social media in medical school	47	8.0	7.0-9.0	.03
No instruction on social media in medical school	22	6.0	4.9-7.4	

[a]Student's t-test, with unequal variance

[b]The possible correct score range was 0-15

Table 2. Added effect of medical school instruction on social media and familiarity with institutional social media policy.

Group	Number	Correct answers mean[b]	P-value[a]
No instruction in medical school, not familiar with policy	18	3.5	.03
Instruction on social media in medical school *only* (31), or familiar with social media policy *only* (4)	35	7.5	
Instruction on social media in medical school and familiar with social media policy	16	9.4	

[a]By Kruskal–Wallis test

[b]The possible correct score range was 0-15

Discussion

The results of this questionnaire reveal several areas of educational opportunity and focus: (1) online privacy and preserving professional and personal identities, (2) maintaining digital professionalism, (3) safeguarding the protected health information of patients, and (4) the impact of existing social media policies and the need for improved education about social media in medicine.

1. Online Security and the Maintenance of Professional Persona versus Personal Identity

As 98% of the resident participants belong to the "millennial" generation, the age range of our subjects is consistent with the majority of medical professionals currently entering clinical practice. Social media is an integral part of daily life for many people in this first generation of "digital natives" and it displays very high levels of connectivity. Over 50% of the participants reported that they are either "almost always" or "mostly" online

and connected. Furthermore, 80% of these digital natives "don't worry at all" or "worry a little" about online privacy [10]. These statistics frame a scenario for trouble when highly connected, digital native users enter the medical arena, where privacy and security are professional and institutional priorities. Social media policy and best practices emphasize the importance of maintaining a distinction between personal and professional online presence.

Some of these advocate a "dual-citizen approach," which maintains online professional and private identities by creating separate online profiles [11]. A majority of medical residents and fellows surveyed was found to have personal information on their profile page [12]. Another study of recent medical graduates revealed that the religious views, sexual orientation, and relationship status could be determined for those who did not activate the privacy settings on their profile [13]. A minority (25%) of our study participants chose "I don't know" as the answer when asked about the effect of privacy settings on a stranger's ability to search for them. Of the subjects considered,

53% (30/57) believed that it is "always okay" to interact with a physician colleague on social media, 49% (28/57) selected "it depends," and 25% (14/57) selected "always okay" to accept a "friend" request from a nurse with whom there is a fleeting professional relationship and no social relationship. This type of online relationship is discouraged by the institution's best practices guidelines. Millennial health care providers appear to have a relaxed stance toward interprofessional digital networking and may not recognize the potential ramifications of blurring their identities online. This suggests an opportunity to inform and counsel during medical school and/or during the transition into clinical practice. On addition of a new 2-hour session called "Online Social Media and Professionalism" to an existing medical school curriculum, 40% of subjects who participated in a 4-month follow-up survey reported that they had "reviewed, edited, and/or made a significant change in their online presence" [14].

Many institutions and policies discourage digital communication with patients. Communicating with patients via an open-access, unsecure site renders protected health information to be vulnerable [15]. Furthermore, interaction with patients via social media blurs the line between personal friend and health care provider. In a survey of medical residents and fellows abroad, approximately half of the respondents felt that "the doctor-patient relationship would be altered if patients discovered that their doctor had a Facebook account" [12]. Interestingly, 33% of our study subjects answered "it depends" when asked about the appropriateness of interacting with patients on social media. A vast majority (89%) of study participants agreed that providing medical advice to patients via social media is "never okay."

2. Digital Professionalism and Institutional Representation

As an employee of a hospital, a physician is a representative of that facility and an ambassador of the medical profession. Furthermore, all physicians are held to local and national standards of digital conduct. For example, the Federation of State Medical Boards has policy guidelines for the appropriate use of social media in medical practice [16]. Activity in the social media world creates new, and possibly unrecognized, responsibilities. Online misadventures, intentional or inadvertent, can result in the breach of patient confidentiality, the fermenting of negative public perception, the undermining of institutional integrity, and the deterioration of professionalism. In a study of recent medical graduates, 46% of subjects had posted their pictures using alcohol and 10% showed intoxication [13]. A study of surgical residents' Facebook posting habits found that 14% of them had posted potentially unprofessional content and 12% had posted clearly unprofessional content [7]. In this study, 49% of participants selected "it depends" while posting a picture of a social event where people are "holding alcoholic beverages." About 26% of the participants stated that they had posted statements or photos on social media that could be considered unprofessional.

A study at a US medical school revealed differences in opinion among students about "appropriate" online content. Some students "felt nothing was inappropriate" and a minority recognized that online activity can reflect on the institution or their profession [17]. A higher incidence of clearly unprofessional Facebook content among surgical residents (12%) compared with surgical faculty (5%) has been reported. Interestingly, among the faculty, men with less than 5 years of clinical practice were associated with unprofessional online habits [6,7]. More conservative attitudes about the appropriateness of Facebook posts have been described among faculty, women, and older participants [18], further emphasizing generational differences in attitudes and habits toward online professionalism. This suggests that younger users see online accountability from a perspective of personal risk only. This validates the need for instruction about the broader impacts of online behavior.

3. Patient Confidentiality

Patient confidentiality achieves wide consensus in the social media discussion. Protected health information and patient privacy should never be compromised by social media or any other form of communication. However, some examples of confidentiality breaches in medicine involve an unwitting rule breaker whose harmful actions are unintentional. For example, a physician was fined by the state medical board and fired by her hospital after posting a detailed account of a patient encounter [3]. Other events described in the media suggest poorer judgment and malicious intent, for example, providers have posted unflattering or personal information about patients on social media sites [5] and have posted inappropriate pictures on image-messaging accounts [4]. Whether intentional or unintentional, the potential for compromising patient information on social media is real and the stakes are quite high.

In this study, subjects were asked about the appropriateness of posting a comment (including a description of gender, approximate age, and injury pattern) about a patient encounter immediately after a mass casualty event. A surprising 25% of respondents selected "always," "it depends," or "not sure." Rarity of disease or injury and proximity in timing or geographic location make identification of the patient in this scenario feasible. Therefore, posting such information could result in the unintentional release of protected health information and, as such, is forbidden at the study institution. This reveals an opportunity to alert our physicians about the hazards to patient privacy in posting seemingly anonymous or harmless information.

4. Social Media Policy and Education

Our results suggest a positive effect from social media instruction and policy familiarity on participants' answer selection in the survey. Although the impact upon online *conduct* of these participants is unknown, improved knowledge and awareness on this topic is desirable. Conventional approaches to professionalism taught in a well-established educational framework may fall short when applied to novel online scenarios in professional medicine, for example, despite reported medical school instruction, 32.5% of interns in this study group answered incorrectly about interacting with patients online. Is this the result of ineffective education or poor knowledge retention? The authors surmise that it is a result of challenges with real-world application at the user level. Traditional

professionalism required during observable human interactions is easily recognized by the learner and has a finite duration. In contrast, social media has the power to project personal activities and musings into the public sphere to lay consumption with infinite duration, even if the user has no such intention. A new approach is needed to address the phenomenon of professionalism in the digital age, or "e-professionalism" [19]. Instruction on "e-professionalism" must provide contextual relevance for the learner and instill an appreciation for the potential reach and impact of online activity.

Bedside Distraction

Another area of digital professionalism not yet adequately studied is the impact of digital networking on clinical performance in the real world. The ubiquity and accessibility of mobile connectivity allows online communication to thrust its way into the clinical environment and, at times, to the bedside. As such, digital professionalism necessarily applies to the clinical workspace. When participants were asked whether logging onto a personal social media account while on duty in a patient care area was acceptable, 27% (17/64) selected "it depends," while 5% (3/64) were "unsure." The social media policy at the study site prohibits this type of activity. When participants were asked whether receiving message notifications on a mobile during work hours in the hospital constitutes a patient safety risk, 9% (5/57)of subjects characterized this practice as "no risk," while 70% (40/57) selected "possible risk." Most would probably contend that receiving audible notifications during patient care activities constitutes, at minimum, a distraction. The presence of distractions and the impact of interruptions on medical care are well documented [20-24]. A case commentary of online messaging distraction resulting in significant medical error has been reported [25]. The impact of messaging alerts and social media activity on physician focus, task performance, and patient safety is an area that requires further investigation.

Looking Ahead

While establishing social media policies, medical institutions must balance several variables; the needs of their professional employees, the potential for social media to improve health care, and their legal and ethical obligations to the patients and communities they serve [26]. Given the high stakes nature of this new digital phenomenon, education has been geared toward establishing acceptable behavioral parameters and has been

rooted in risk prevention [2,14,15]. Some have raised concerns that an overemphasis on control, security, and risk avoidance may impede the potential benefits that social media can provide [27].

A sophisticated understanding and awareness of the potential benefits and hazards of social media are a prerequisite to responsible online activity. Efforts to raise awareness among medical students and physicians in transition may help shape desired online behaviors. Evidence supporting this approach includes a report of a "raised perception of risk" among medical students after hearing personal stories, warnings by school officials, and media reports [17]. While the debate about effective and comprehensive instruction for online users continues, there are resources available that provide current and relevant information about judicious use of social media in the context of medical practice [15,28,29].

Limitations

Some subjects were previously enrolled as students at the study institution so their "formal instruction" on social media usage may have contributed to familiarity with the current institutional policy. We did not identify those individual subjects. The study is limited by a relatively small sample size (n=124 invited subjects) at a single academic medical institution. Questionnaires were delivered electronically via their respective program director during a very busy onboarding process, which may contribute to a response bias. Subjects may have felt compelled to complete the survey or may not have given proper attention to each question, given all the activity that goes on during the onboarding process. Only incoming house officers were surveyed which may make some of the conclusions less generalizable to all house officers.

Conclusion

Young physicians demonstrate a casual approach to social media activity in the context of professional medical practice. However, social media instruction and/or familiarity with the social media policy are associated with more cautious perceptions about online behavior. Furthermore, assessment of perceptions and practices of new employees in a health care environment may help improve the content and delivery of policy information. This approach may help to preserve institutional priorities, protect patient information, and safeguard young professionals from online misadventures.

Conflicts of Interest

None declared.

Multimedia Appendix 1

[PDF File (Adobe PDF File), 32KB - jmir_v18i6e119_app1.pdf]

References

1. Hamm MP, Chisholm A, Shulhan J, Milne A, Scott SD, Klassen TP, et al. Social media use by health care professionals and trainees: a scoping review. Acad Med 2013 Sep;88(9):1376-1383. [doi: 10.1097/ACM.0b013e31829eb91c] [Medline: 23887004]

2.   Kind T, Gengrich G, Sodhi A. Medical Education Online. 2010 Sep 15. Social Media Policies at US Medical Schools URL: http://www.webcitation.org/6f68E83rE [accessed 2015-10-02] [WebCite Cache ID 6f68E83rE]

3.   Conaboy C. Boston.com. 2011 Apr 20. For doctors, social media a trick case URL: http://www.webcitation.org/6f68hK3a3 [accessed 2015-09-28] [WebCite Cache ID 6f68hK3a3]

4.   Neporent L. ABC news. 2014 Jul 08. Nurse firing highlights hazards of social media in hospitals URL: http://abcnews.go.com/Health/nurse-firing-highlights-hazards-social-media-hospitals/story?id=24454611 [accessed 2015-09-28] [WebCite Cache ID 6f68pymxb]

5.   Kinsey M. Slate.com. 2014 Jan 10. What happens in the hospital doesn't stay in the hospital URL: http://www.webcitation.org/6f6IcIYVp [accessed 2015-09-28] [WebCite Cache ID 6f6IcIYVp]

6.   Langenfeld SJ, Sudbeck C, Luers T, Adamson P, Cook G, Schenarts PJ. The Glass Houses of Attending Surgeons: An Assessment of Unprofessional Behavior on Facebook Among Practicing Surgeons. J Surg Educ 2015;72(6):e280-e285. [doi: 10.1016/j.jsurg.2015.07.007] [Medline: 26276300]

7.   Langenfeld SJ, Cook G, Sudbeck C, Luers T, Schenarts PJ. An assessment of unprofessional behavior among surgical residents on Facebook: a warning of the dangers of social media. J Surg Educ 2014;71(6):e28-e32. [doi: 10.1016/j.jsurg.2014.05.013] [Medline: 24981657]

8.   Greysen SR, Chretien KC, Kind T, Young A, Gross CP. Physician violations of online professionalism and disciplinary actions: a national survey of state medical boards. JAMA 2012 Mar 21;307(11):1141-1142. [doi: 10.1001/jama.2012.330] [Medline: 22436951]

9.   Chretien KC, Greysen SR, Chretien J, Kind T. Online posting of unprofessional content by medical students. JAMA 2009 Sep 23;302(12):1309-1315. [doi: 10.1001/jama.2009.1387] [Medline: 19773566]

10.   Rosenstiel T, Sonderman J, Loker K. American Press Institute.: American Press Insitute; 2015 Mar 16. Digital Lives of Millennials URL: http://www.webcitation.org/6f6InTqaa [accessed 2015-10-01] [WebCite Cache ID 6f6InTqaa]

11.   Mostaghimi A, Crotty BH. Professionalism in the digital age. Ann Intern Med 2011 Apr 19;154(8):560-562. [doi: 10.7326/0003-4819-154-8-201104190-00008] [Medline: 21502653]

12.   Moubarak G, Guiot A, Benhamou Y, Benhamou A, Hariri S. Facebook activity of residents and fellows and its impact on the doctor-patient relationship. J Med Ethics 2011 Feb;37(2):101-104. [doi: 10.1136/jme.2010.036293] [Medline: 21160080]

13.   MacDonald J, Sohn S, Ellis P. Privacy, professionalism and Facebook: a dilemma for young doctors. Medical Education 2010;44(8):805-813. [doi: 10.1111/j.1365-2923.2010.03720.x]

14.   Lie D, Trial J, Schaff P, Wallace R, Elliott D. "Being the best we can be": medical students' reflections on physician responsibility in the social media era. Acad Med 2013 Feb;88(2):240-245. [doi: 10.1097/ACM.0b013e31827bf5fc] [Medline: 23269290]

15.   Chretien KC, Kind T. Social media and clinical care: ethical, professional, and social implications. Circulation 2013 Apr 2;127(13):1413-1421 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.112.128017] [Medline: 23547180]

16.   Rhyne J, Talmage L, Kopetski J. Federation of State Medical Boards. Model Policy Guidelines for the Appropriate Use of Social Media and Social Networking in Medical Practice URL: https://www.fsmb.org/Media/Default/PDF/Publications/pub-social-media-guidelines.pdf [accessed 2016-04-28] [WebCite Cache ID 6h73Mx7YS]

17.   Chretien KC, Goldman EF, Beckman L, Kind T. It's your own risk: medical students' perspectives on online professionalism. Acad Med 2010 Oct;85(10 Suppl):S68-S71. [doi: 10.1097/ACM.0b013e3181ed4778] [Medline: 20881708]

18.   Jain A, Petty EM, Jaber RM, Tackett S, Purkiss J, Fitzgerald J, et al. What is appropriate to post on social media? Ratings from students, faculty members and the public. Med Educ 2014 Feb;48(2):157-169. [doi: 10.1111/medu.12282] [Medline: 24528398]

19.   Cain J, Romanelli F. E-Professionalism: A New Paradigm for a Digital Age. Curr Pharm Teach Learn 2009;1:66-70.

20.   Hedberg B, Larsson US. Environmental elements affecting the decision-making process in nursing practice. J Clin Nurs 2004 Mar;13(3):316-324. [Medline: 15009334]

21.   Rhoades DR, McFarland KF, Finch WH, Johnson AO. Speaking and interruptions during primary care office visits. Fam Med 2001;33(7):528-532. [Medline: 11456245]

22.   Chisolm C, Collison E, Nelson D. Emergency department workplace interruptions: are emergency physicians "interrupt-driven" and "multitasking"? Acad Emerg Med 2000;7:1239-1243.

23.   Flynn EA, Barker KN, Gibson JT, Pearson RE, Berger BA, Smith LA. Impact of interruptions and distractions on dispensing errors in an ambulatory care pharmacy. Am J Health Syst Pharm 1999 Jul 1;56(13):1319-1325. [Medline: 10683129]

24.   Brixley J, Tang Z, Robinson D. Interruptions in a level one trauma center: a case study. Int J Med Inf 2008;77:235-241.

25.   Halamka J. U.S. Department of Health and Human Services. Agency for Healthcare Research and Quality; Patient Safety Network. 2011 Dec. Order Interrupted by Text: Multitasking Mishap URL: https://psnet.ahrq.gov/webmm/case/257 [accessed 2015-12-31] [WebCite Cache ID 6f6J7FbCe]

26.   Chretien KC, Kind T. Climbing social media in medicine's hierarchy of needs. Acad Med 2014 Oct;89(10):1318-1320 [FREE Full text] [doi: 10.1097/ACM.000000000000430] [Medline: 25076202]

27.   Fenwick T. Social media and medical professionalism: rethinking the debate and the way forward. Acad Med 2014 Oct;89(10):1331-1334 [FREE Full text] [doi: 10.1097/ACM.000000000000436] [Medline: 25076200]

28.   Fogelson NS, Rubin ZA, Ault KA. Beyond likes and tweets: an in-depth look at the physician social media landscape. Clin
       Obstet Gynecol 2013 Sep;56(3):495-508. [doi: 10.1097/GRF.0b013e31829e7638] [Medline: 23835911]

29.   Bernhardt JM, Alber J, Gold RS. A social media primer for professionals: digital dos and don'ts. Health Promot Pract 2014
       Mar;15(2):168-172. [doi: 10.1177/1524839913517235] [Medline: 24396119]

XSL•FO
RenderX

Original Paper

Demographic-Based Content Analysis of Web-Based Health-Related Social Media

Shouq A Sadah[1], MS; Moloud Shahbazi[1], MS; Matthew T Wiley[1,2], MS; Vagelis Hristidis[1], PhD

[1]University of California, Riverside, Department of Computer Science and Engineering, Riverside, CA, United States
[2]SmartDocFinder LLC, Riverside, CA, United States

Corresponding Author:
Shouq A Sadah, MS
University of California, Riverside
Department of Computer Science and Engineering
351 Winston Chung Hall
900 University Ave
Riverside, CA, 92521
United States
Phone: 1 9517565691
Fax: 1 9518274643
Email: ssada003@cs.ucr.edu

Abstract

Background: An increasing number of patients from diverse demographic groups share and search for health-related information on Web-based social media. However, little is known about the content of the posted information with respect to the users' demographics.

Objective: The aims of this study were to analyze the content of Web-based health-related social media based on users' demographics to identify which health topics are discussed in which social media by which demographic groups and to help guide educational and research activities.

Methods: We analyze 3 different types of health-related social media: (1) general Web-based social networks Twitter and Google+; (2) drug review websites; and (3) health Web forums, with a total of about 6 million users and 20 million posts. We analyzed the content of these posts based on the demographic group of their authors, in terms of sentiment and emotion, top distinctive terms, and top medical concepts.

Results: The results of this study are: (1) Pregnancy is the dominant topic for female users in drug review websites and health Web forums, whereas for male users, it is cardiac problems, HIV, and back pain, but this is not the case for Twitter; (2) younger users (0-17 years) mainly talk about attention-deficit hyperactivity disorder (ADHD) and depression-related drugs, users aged 35-44 years discuss about multiple sclerosis (MS) drugs, and middle-aged users (45-64 years) talk about alcohol and smoking; (3) users from the Northeast United States talk about physical disorders, whereas users from the West United States talk about mental disorders and addictive behaviors; (4) Users with higher writing level express less anger in their posts.

Conclusion: We studied the popular topics and the sentiment based on users' demographics in Web-based health-related social media. Our results provide valuable information, which can help create targeted and effective educational campaigns and guide experts to reach the right users on Web-based social chatter.

Introduction

As Web-based social media are growing in popularity, the number of people who share their experiences or ask for support in health-related social media has also increased [1]. Fox and Jones have found that 41% of e-patients have read someone else's commentary or experience about health on a Web-based news group, website, or blog [2]. Kane et al [3] reported that more than 60 million Americans read or contribute to Health 2.0 apps, in which they consider these apps as their first source

when gathering data and opinions. About 40% of Americans doubt a professional opinion when it conflicted with what they form from Web-based health social media [3].

One of the key benefits of health-related Web-based social media reported by researchers is the increased access to information to various demographic groups, regardless of age, education, income, or location [4]. However, previous work has mainly relied on user surveys to study the effect of the use of social media to health-related factors such as psychological distress [5]. In addition, previous work does not reveal granular information on what disorders or other health topics are mostly discussed in the Internet by each demographic group, which would allow health care providers to create targeted and effective educational campaigns.

In this work, we conducted the first, to our best knowledge, large-scale data-driven comparative analysis of the content of health-related social media across various demographic dimensions—gender, age, ethnicity, location, and writing level. For each demographic group, we study the content of the posts across the following dimensions: sentiment, popular terms (keywords), and medical concepts (particularly disorders and drugs). Concepts refer to entries in the Unified Medical Language System (UMLS) vocabulary [6], whereas terms are just words from the posts' text that may or may not belong to any UMLS concept. We report results for 3 types of social media: (1) general Web-Based Social Networks, namely Google+ and Twitter, (2) drug review websites, and (3) health Web forums. The selection of social media types was based on their popularity and on our study of the literature on health-related social content [7]. The objective of this study was to identify which health topics are discussed in which social media by which demographic groups, to better guide educational outreach and research activities.

Related Work

Analysis of Health-Related Social Media

Different studies were established and conducted by researchers to study the effectiveness of Web-based social media in changing and improving the communication between providers and patients. Hackworth and Kunz [8] reported that 80% of Americans have searched the Internet for health-related information. Grajales et al [9] illustrated how, when, and why social media are used by health care sectors by conducting a narrative review of case studies, and they provided 4 recommendations that stakeholders may consider to engage with social media. Because analyzing the health-related content of social media is increased recently [10], Denecke and Nejdl [11] performed content analysis of medical concepts in different health-related social media sources. They presented a method to classify posts as informative or affective, and they found that doctors share health-related information, whereas patient and nurses are more likely to share personal experiences. Lu et al [12] analyzed the content of 3 disease-specific health communities including lung cancer, breast cancer, and diabetes and defined their relationship to 5 main informative topics: symptoms, complications, examination, drugs, and procedures. This study shows that examination is a hot topic for users with breast cancer, whereas symptoms are more likely to be discussed

by users with lung cancer. Wiley et al [13] analyzed the content of drug-related chatter on various social media forums. The study demonstrates that Web-based social media's characteristics such as moderation affect the discussions in different ways including subjectivity and type of drugs discussed.

Measuring and Estimating Demographics of Users of Social Media

Krueger et al [14] studied the mortality attributable to low education level in the United States. They found people with less than high school degree have more mortality rate; thus, improving the US educational attainment could increase the survival in US population. A Pew research conducted in 2012 showed that white ethnicity represents 75% of social media websites users, where women in age group of 30-49 years participate more in these websites [15]. Another study by eMarkter found that Hispanic are more active in social media with 68.9% of them using social networks compared with 66.2% of total US population [16]. Mislove et al [17] estimated gender and ethnicity for Twitter users. The gender is estimated by using the reported first name and comparing it to the 1000 most popular first names reported by the US Social Security Administration, whereas ethnicity is estimated by using the reported last name and comparing it to the frequently occurring surnames reported by 2000 US Census. Using Mislove's gender classifier, Mandel et al [18] analyzed the tweets related to Hurricane Irene. Liu et al [19] proposed Natural Language Processing (NLP) methods to extract the demographics (gender, age, ethnicity) of users of social posts. Anderson-Bill et al [20] recruited Web-health users to examine their demographics, behavioral, and psychosocial characteristics, and they found that Web-health users are more likely middle-aged, upper class, and well-educated women. Although the aforementioned work examined health-related social media and their content, none of them studied how different demographics use Web-based social media, which is studied in this work.

Sadah et al [21] studied how many users from each demographic group (by gender, age, ethnicity, location, and writing level) participate in various social media, but it did not study the content of the posts, which is the focus of this paper. Some of the key findings of that work are: (1) drug review websites and health Web forums are dominated by female users; (2) the participants of health-related social media are generally older with the exception of the 65+ years bracket; (3) Asian and black ethnic groups are underrepresented in drug review websites and health Web forums, and blacks are also underrepresented in health-related Web-based social networks; (4) users in areas with better access to health care participate more in Web-based health social media; and (5) the writing level of users in health social media is significantly lower than the reading level of the population.

Methods

Key Challenges

A key challenge is to estimate the demographic group, for example, gender, of a Web-based user when this information is not explicitly stated. Another challenge in this work is the

extraction of medical concepts from social posts, given that existing tools such as MetaMap focus on biomedical text, which is generally generated by researchers or practitioners; therefore, we filtered out some misclassified concepts generated by this tool to work on health social media posts. Another challenge has been the time to extract the medical concepts from the social posts. In this paper, we process more than 20 million posts, which would take several months to parse on a single machine. For that, we have parallelized this into 10 machines that extracted all concepts in about 1 month. To extract popular terms for each demographic group, we use stemming to merge together terms with the similar root.

Datasets

As summarized in Table 1, for general social networks, we chose Twitter and Google+ for their popularity and number of users (we did not include Facebook as it does not provide public data). For the other 2 types, we selected 3 different websites for each one to ensure diversity. More information about the sources including start and end date is available in Table A.1 and A.2 of Multimedia Appendix 1. Because Twitter and Google+ are

general social networks, we filtered the posts using 276 representative health-related keywords as follows: (1) Drugs: from the most prescriptions dispensed from RxList.com, we selected the 200 most popular drugs [22]. By removing the variants of the same drug (eg, different milligram dosages), the final list of drugs contained 125 unique drug names. (2) Hashtags: from Twitter Hashtags, we selected 11 popular health-related Twitter hashtags such as #BCSM (Breast Cancer and Social Media). (3) Disorders: 81 popular disorders were selected such as AIDS and asthma. (4) Pharmaceuticals: the 12 largest pharmaceutical companies were selected such as Novartis. (5) Insurance: 44 of the biggest insurances were selected such as Aetna and Shield. The rationale of selecting the keywords was to cover as much as we can by including popular drugs and disorders, popular health-related hashtags in Twitter, and other related health keywords that can help increase the number of the posts related to health, similar to previous work on Twitter filtering [13,21]. A complete list of used keywords can be found in Table B.1 of Multimedia Appendix 1, and all terms' frequencies for both sources can be found in Table B.2 of Multimedia Appendix 1.

Table 1. List of all used sources with their number of posts and with the available demographic attributes.

Dataset	No. of posts	Gender[a]	Age[a]	Ethnicity[a]	Location[a]	Writing level
TwitterHealth [23]	11,637,888	Gender classifier	NO	Ethnicity classifier	YES	Writing level classifier
Google+Health [24]	186,666	YES	YES	Ethnicity classifier	YES	Writing level classifier
Drugs.com [25]	74,461	Gender classifier	NO	NO	NO	Writing level classifier
DailyStrength/Treatments [26]	1,055,603	YES	YES	NO	YES	Writing level classifier
WebMD/Drugs [27]	122,040	YES	YES	NO	NO	Writing level classifier
Drugs.com/Answers [28]	320,118	Gender classifier	NO	NO	NO	Writing level classifier
DailyStrength/Forums [29]	5,948,877	YES	YES	NO	YES	Writing level classifier
WebMD [30]	1,128,629	Gender classifier	NO	NO	NO	Writing level classifier

[a]*NO* indicates that the demographic attribute is not provided by the source and no classifier is used due to low accuracy. *YES* indicates that the demographic attribute is provided by the source. More details on the demographic classifiers are available in the paper by Sadah et al [21].

Then, to filter out Twitter using the health-related keyword list, we used the Twitter streaming Application Program Interface (API) [31] to extract the relevant tweets for TwitterHealth. Google+Health posts were collected via the Google+ API [32], in which the health-related keyword list was used in the queries to obtain relevant posts for Google+. For the other drug review websites and health Web forums, we built a crawler for each website in Java using the Java library jsoup [33] for extracting

and parsing hypertext markup language content. For each website, we crawled and collected the available data, including public user information, posts, disorders, conditions, keywords, tags, rating, and so forth. We emphasize that we do not collect or use any private data, and we only collected publicly available data in accordance with each site's terms of use. Figure 1 shows the overall process of our analysis.

XSL·FO

RenderX

Figure 1. Overview of the data collection and analysis process.
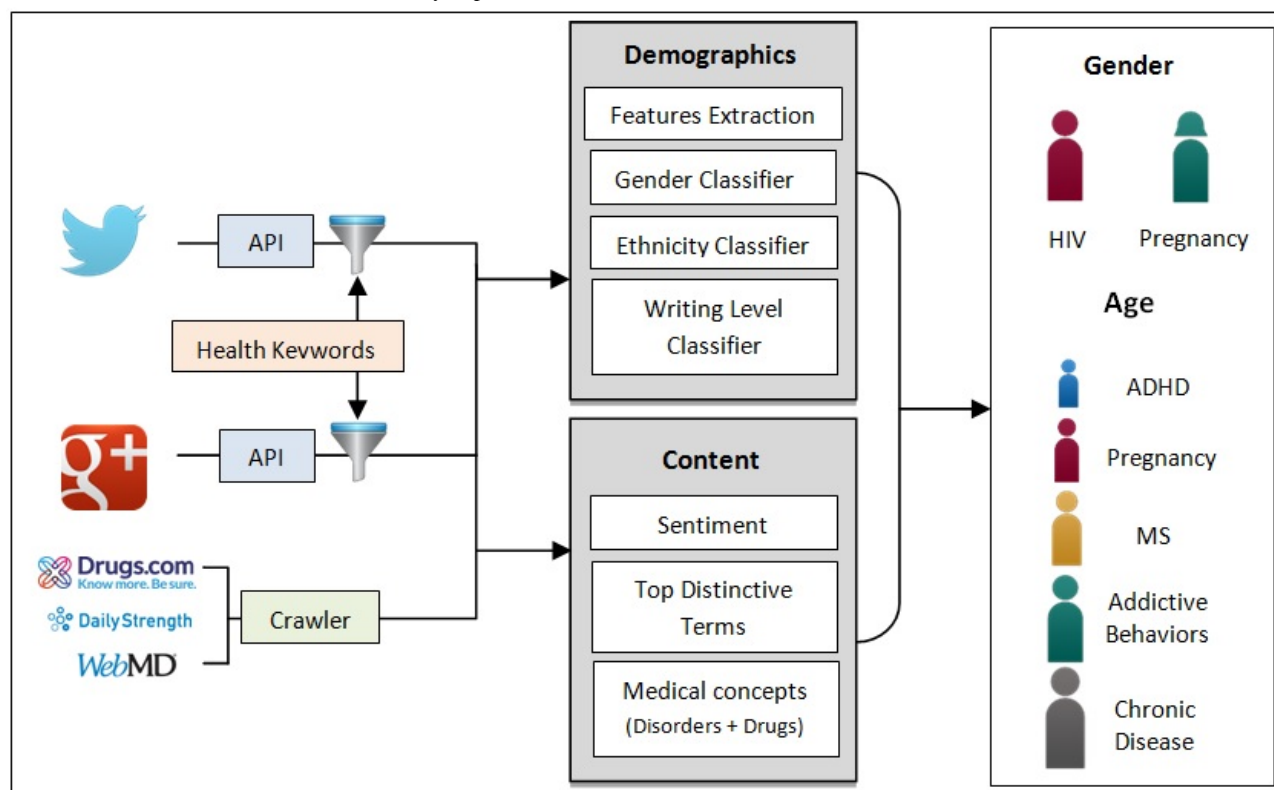
Demographic Data Computation

The demographic data (age, gender, ethnicity, location, writing level) of users are extracted from the data or estimated through classifiers as discussed in the study by Sadah et al where more statistics of the collected posts such as dates and number of users are also reported [21]. As summarized in Table 1, gender attribute is either reported by the source or generated by a classifier that uses the first name to distinguish between male and female. Age and location, on the other hand, are used as they reported by the source; however, location was further processed to map user's input into geolocations using Google API [32]. Because ethnicity is not reported in any source, we used a classifier that uses the last name to predict ethnicity for users in sources that provide the last name. For writing level, we measured each user writing level using a modified version of Flesch–Kincaid Grade Level [34].

Sentiment and Emotion

To compute sentiment and emotion, we map each phrase in the post to a phrase from a sentiment lexicon. We use a sentiment lexicon, SentiWordNet [35], and an emotion lexicon, NRC word-emotion lexicon [36]. These 2 lexicons were selected owing to their effectiveness and popularity in previous studies [11,37,38] and because they cover complementary aspects. We use the SentiWordNet dictionary for sentiment, which assigns positive, negative, and objective score to each term where the sum of all 3 scores equals 1. Because SentiWordNet uses senses and part of speech, the Stanford CoreNLP Trigger [39] was used to tag each word with its part of speech tag. All words in posts and SentiWordNet were then stemmed to remove words variation. The longest possible match is then used to map each phrase in posts to a phrase from SentiWordNet, and after that,

each post's sentiment is calculated by averaging scores of all phrases. For each source, the total sentiment score for each demographic attribute is measured by averaging all posts' scores associated with that attribute and normalized by the number of posts of the attribute. For emotion, we use the NRC word-emotion lexicon, which measures anger–fear, trust–disgust, and anticipation–surprise.

Top Distinctive Terms

The content of all sources was analyzed to get the top distinctive terms for each source. All posts are first filtered to remove stop words and then stemmed using Porter stemmer [40]. From these words, we considered only the ones that occur in at least 0.01% of the total number of posts that are annotated for a given demographic attribute value (or 30 if 0.01% is less than 30). That is, if less than 0.01% of posts from users who reported their gender contain a term, this term is not reported in either male or female group analysis. Then, for each demographic attribute value, that is, male, we normalized the number of occurrences for each term in that attribute value by the number of users posts in the same attribute value to get the frequency, for example:

Freqmale (headache)=No. of occurrences (headache) in male/No. of male posts [1]

To get the top 10 distinctive terms for each demographic attribute, we then calculated the relative difference as follows:

RelDifmale (headache)=[Freqmale (headache) – AvgFreqgender (headache)] / AvgFreqgender (headache) [2]

Where $AvgFreq_{gender}$ (*headache*) is the average frequency of the word headache in all posts by male or female users. For

example, $AvgFreq_{location}(headache) = [Freq_{Northeast}(headache) + Freq_{Midwest}(headache) + Freq_{South}(headache) + Freq_{West}(headache)] /4$. Finally, we only display health-related terms in each demographic group that have a relative difference greater than 0.1; that is, we decided to hide results with a difference of less than 10% from the average score, which we believe is intuitive.

Medical Concepts

To annotate posts with corresponding medical concepts from the UMLS [41], the MetaMap tool [42] was used to represent each post as a set of medical concepts.

Because MetaMap was originally built to extract concepts from biomedical text generated by researchers or practitioners, it is not perfect to annotate social media posts [43]. Therefore, we manually removed some annotations that were misclassified by MetaMap as following: (1) we order generated concepts by their frequencies for each source systematically, (2) we analyze each phrase that was mapped for each concept, and (3) we delete the misclassified UMLS concepts from the results. For example, the letter "i" mapped to (immunologic factor) and word bad mapped to (organic chemical). Such mistakes were deleted from MetaMap annotations to improve accuracy. In UMLS, we have 15 semantic groups (eg, Disease or Anatomy), and each concept in UMLS is associated with one or more semantic types, where each semantic type belongs to 1 semantic group. In this part, we analyzed only 2 semantic groups including drugs and disorders, and we reported the top distinctive drugs and disorders for each demographics using the same threshold and method used in finding top distinctive terms (Equation 2).

Results

In this section, we present our results for sentiment and emotion, top distinctive terms, and medical concepts by each demographic group. Two medical concept types were considered and reported to avoid less interesting results: disorders and drugs. For each demographic group, we show the top distinctive disorders and drugs using Equation 2 that have a relative difference more than 0.1. Some demographic attribute values are not reported owing to small number of users (age group (0-17) and (65+) in Google+Health), or demographic attribute is not reported by the source (all age groups in TwitterHealth), or because users talk about unrelated health topics (writing level (0-5) in TwitterHealth talk about astrology), or the relative difference (Equation 2) for the top findings is less than 0.1.

Gender

In Table 2, we summarize the top distinctive (highest relative difference according to Equation 2) terms by gender; note that some demographic attributes such as female in Google+Health do not have distinctive terms. Because Twitter and Google+ are more news-based social media, many health posts share news in different areas including politics and sports—we excluded them to include health-related keywords only. Our first key finding is that male users in TwitterHealth tend to talk more about the reproductive system, tumor and AIDS, and health insurance, whereas female users talk about headache and emotion. In drug review websites and health Web forums, female users tend to talk more about pregnancy-related topics, whereas male users discuss pain drugs, cholesterol, and heart problems.

Table 2. Top 10 distinctive terms by gender.

Gender	TwitterHealth	Google+Health	Drugs	Forums
Male	Prostate, Gay, Testicular, Viagra, Tumor, AIDS, Obamacare, Marijuana, Medicare, Insurance	Pharmacology, Encephalomyelitis, Amphetamine, Pertussis, Fukushima, Pfizer, Novartis, Neutrophil, Biomed, Viagra	Wife, Oxycontin, Urine, Lisinopril, Cholesterol, Hydrocodone, Disc, Spinal, Libido, Diovan	HIV, Wife, Tinnitus, Gay, Cholesterol, Artery, AA (alcoholic anonymous), Valium, Cardiologist, Alcohol
Female	Cry, Migraine, Moody, Frown, Pound, Laugh, Nap, Eczema, Headache, Tension	N/A	Ovulation, IUI (intrauterine insemination), Pregnancy, Clomid (used to cause ovulation in women), IVF (in vitro fertilization), Pregnant, Birth, Boyfriend, BC (birth control), Fibromyalgia	Miscarry, PCOS (polycystic ovary syndrome), Endometriosis, Lupron, Uterus, Hysterectomy, Infertility, Ovarian, Rheumatologist, Progesterone

In Table 3, we summarize top distinctive disorders by gender. Male users in drug review websites mainly talk about back pain and blood pressure, whereas female users talk about pregnancy. In health Web forums and websites, male users discuss heart problems and panic topics, and female users talk more about skin disorders, headache, and chronic fatigue disorders. In TwitterHealth and Google+Health, top disorders discussed by male users can be classified as sexually transmitted diseases, including AIDS and herpes, whereas female in TwitterHealth as seen in the top distinctive terms discuss topics related to headache and feelings.

Table 3. Top 5 distinctive disorders by gender.

Gender	TwitterHealth	Google+Health	Drugs	Forums
Male	Acquired Immunodeficiency Syndrome (AIDS), HIV seropositivity, Cerebrovascular accident (stroke), Incised wound, Herpes NOS	Gonorrhea, Marijuana abuse, Sexually transmitted diseases, Malignant neoplasm of lung, Infantile neuroaxonal dystrophy	Low back pain, Dry cough, Blood pressure finding, Back pain, Diabetic	Atrial fibrillation, Codependency, Panic attacks, Diabetes, Marijuana abuse
Female	Migraine disorders, Emotional, Headache, Pain NOS adverse event, Asleep	Chronic fatigue syndrome	Gravidity; Endometriosis, site unspecified; Yeast infection; Fibromyalgia; Hot flushes	Dermatitis herpetiformis, familial; Lupus vulgaris; Lupus erythematosus, systemic; Fibromyalgia; Migraine disorders

Table 4 summarizes top distinctive drugs by gender. In drug review websites, the top drugs discussed by female users are related to pregnancy including birth control and ovulation stimulation, whereas male users talk mainly about drugs related to blood pressure. In health Web forums ,male users discuss depression-related drugs and alcohol topics. In TwitterHealth, not many distinctive drugs were found for female and male users, whereas in Google+Health, different drugs and chemicals were reported.

Sentiment and emotion were evaluated for all sources. Because the results look similar between gender groups, we summarize the results in Tables C.1 and C.2 of Multimedia Appendix 1.

Table 4. Top 5 distinctive drugs by gender.

Gender	TwitterHealth	Google+Health[a]	Drugs[a]	Forums[a]
Male	Viagra	Aldosterone, DC101 monoclonal antibody, Bicarbonates, Aspartame, Methamphetamine[1]	Low-density lipoproteins, Plavix, Bystolic[6], Oxycodone, Opiates	Alcohols, Xanax[4], Detox adjuvant[2], Prozac[4], Dietary lead
Female	Trivalent influenza vaccine	Thioctic acid, Detoxadjuvant[2], Seroquel	Yaz[3], Implanon[3], Tamoxifen[2], Estrogens, Clomid[3]	Plaquenil, Diamox, Topamax, Concerta, Synthroid

[a]Some of the drugs are coded to match the corresponding disorders they treat:[1]ADHD,[2]Cancer,[3]pregnancy,[4]depression,[5]MS,[6]BP, heart problem and cholesterol,[7]Diabetes.

Age

Table 5 summarizes the top 10 distinctive terms for each age group. Generally, for younger groups (0-17 years), ADHD and skin problems are popular topics in drug review websites, whereas in health Web forums, they talk more about parents and homosexuality. For age groups of 18-34 years in drug review websites and health Web forums, the main topics discussed are related to relationships, pregnancy, or getting pregnant using simple intervention methods, or family members; whereas the same groups in Google+Health talk about different aspects including vitamins and sleep disorders. Age group of 35-45 years also discusses pregnancy topics but using sophisticated intervention methods including *in vitro fertilization*. Age group of 45-64 years, as in Table 4, discusses topics related to chronic diseases including fibromyalgia, disc, and cholesterol, and it also discusses other topics including addiction to smoking, alcohol, and menopause. HIV also appears to be a popular topic for that group in Google+Health and health Web forums. Finally, people aged older than 65 years also talk more about chronic diseases and heart-related problems including drugs that can help mitigate the pain. We see that most topics are more likely discussed by women because drug review websites and health Web forums are dominated by female users [21].

Table 5. Top 10 distinctive terms by age.

Age, years	Google+Health	Drugs	Forums
0-17	N/A	Concerta, Acne, ADHD, Birth, Wash, Lip, Prescribed, Boyfriend, Skin, Scar	Lesbian, Bullying, Buddy, Gay, Mum, Crush, Suicide, Rape, Teen, Dad
18-34	Supplement, Arthritis, Weight, Vitamin, Headache, Hospital, Friend, Food, Love, Skin	Clomid (used to cause ovulation in women), Ovulation, Phentermine, Calorie, Pregnancy, Gym, Pregnant, Baby, BC (birth control), Workout	BC (birth control), Clomid (used to cause ovulation in women), Ovulation, PCOS (polycystic ovary syndrome), TTC (trying to conceive), Miscarried, Fiance, Baby, Pap, Conceive
35-44	Vitamin, Sleep, Food, Parkinson, Friend, Healthcare, Community, Vaccine, Pain, Insomnia	IVF (in vitro fertilization), IUI (intrauterine insemination), Clomid, Ovulation, Marriage, Divorce, Mania, Narcotic, Lithium, Kid	IVF (in vitro fertilization), IUI (intrauterine insemination), BFP (big fat positive), BFN (big fat negative), PG, Stbx (Soon-to-be-ex), Lupron, HCG, Infertility, Fertile
45-64	Syndrome, Death, Chronic, Diet, Anxiety, Hospital, HIV, Infect, Treatment, Flu	Menopause, Fibromyalgia, Oxycontin, Chantix, AA (alcoholic anonymous), RA (rheumatoid arthritis), Disc, Narcotic, Heat, Chronic	Menopause, Grandson, HIV, Disc, Tinnitus, Lesion, Liver, Cholesterol, Enzyme, Colon
65+	N/A	Diovan, Lisinopril, Neuropathic, Urine, Ankle, Cholesterol, Stroke, Arthritis, BP (blood pressure), Cancer	COPD (chronic obstructive pulmonary disease), Valium, PD (panic disorder), Caregiver, Retire, Oxygen, Transplant, Chemo, Cardiologist, Grandchildren

Table 6 summarizes top distinctive disorders by age. In drug review websites, the young age group of 0-17 years talks more about skin disorders and mental disorders, whereas the same age group in health forums websites discusses mainly mental disorders. Age groups of 18-34 years and 35-44 years in drug review websites talk about pregnancy and mental disorder topics. Older age groups in both sources tend to talk about diabetes, heart diseases, and muscles pain.

Table 6. Top 5 distinctive disorders by age.

Age, years	Drugs	Forums
0-17	Acne vulgaris, Acne, Attention deficit hyperactivity disorder, Mood swings, Feeling suicidal (finding)	Depressed mood, Incised wound, Mental depression, Fear (finding), Emotional distress
18-34	Endometriosis, site unspecified, Gravidity, Panic attacks, Anxiety attack, Manic	Gastritis, Asthma, Panic, Anxiety disorders, Observation of attack
35-44	Endometriosis, site unspecified, Manic, Manic mood, Addictive behavior, Chronic pain	Autistic disorder; Disability; Lupus erythematosus, systemic; Attention deficit hyperactivity disorder; Pressure (finding)
45-64	Hot flushes, Chronic pain, Fibromyalgia, Night sweats, Nerve pain	Codependency; Gastritis; Fibromyalgia; Lupus vulgaris; Lupus erythematosus, systemic
65+	Muscle cramps in leg, Dry cough, Lassitude, Diabetic, Blood pressure finding	Atrial fibrillation, Diabetic, Panic attacks, Cerebrovascular accident, Gastroesophageal reflux disease

In Table 7, we summarize all age groups' top drugs. For the younger group of 0-17 years in drug review websites, top drugs discussed are the ones related to ADHD. Age group of 18-34 years in drug review websites discusses pregnancy-related drugs, whereas for age group of 35-44 years, the top drugs are related to MS disorder. This group of 35-44 years in health Web forums tends to share information about ADHD drugs. Older age users (65+ years) discuss drugs related to heart problems, blood pressure, diabetes, and cholesterol.

Table 7. Top 5 distinctive drugs by age.

Age, years	Drugs[a]	Forums[a]
0-17	Accutane, Concerta[1], Vyvanse[1], Strattera[1], Implanon[3]	Commit Lozenge, Relate—vinyl resin, Vent, Zoloft[4], Topamax
18-34	Clomid[3], Phentermine, Seasonique[3], Lupron[2], Yaz[3]	Human papilloma virus vaccine, Topamax, Diamox, Adderall[1], Antibiotics
35-44	Clomid[3], Rebif[5], Avonex[5], Tysabri[5], Lortab	Concerta[1], Melatonin, Diamox, Plaquenil, Adderall[1]
45-64	Tamoxifen[2], Avonex[5], Oxycontin, Savella, Soma	Smoke, Hydrocortisone, Cymbalta[4], Lyrica[7], Alcohols
65+	Plavix[6], Diovan[6], Actos[7], Hydroxymethylglutaryl-CoA reductase inhibitors, Lipitor[6]	Metformin[7], Carbohydrates, Oxygen, Sugars, Xanax[4]

[a]Some of the drugs are coded to match the corresponding disorders they treat: [1]ADHD,[2]Cancer,[3]pregnancy,[4]depression,[5]MS,[6]BP, heart problem and cholesterol,[7]Diabetes.

Sentiment and emotion were evaluated for all sources. Because the results look similar among age groups, we summarize the results in Tables C.3 and C.4 of Multimedia Appendix 1. One key finding from the emotion results is that older people in Google+Health and drug review websites express less anger, whereas younger people in drug review websites express more anger.

Ethnicity

Only TwitterHealth and Google+Health have a large enough number of users whose ethnicity we can estimate (see Table A.2 in Multimedia Appendix 1), and hence, we only report finding for these outlets. In Table 8, we summarize top disorders for each ethnicity except black owing to the small number of users. As a key finding of top disorders, fibromyalgia is one of the top disorders that white and Hispanic users discuss in TwitterHealth, heart and kidney diseases are discussed more by Asian users, and headache and sleeplessness are 2 of the top disorders discussed by Hispanic users. The other ethnicity-based results exhibit less variance among the ethnicity groups, and hence, we report in Tables C.5, C.6, C.7, and C.8 of Multimedia Appendix 1.

Table 8. Top 5 distinctive disorders by ethnicity.

Ethnicity	TwitterHealth	Google+Health
White	Fibromyalgia, Presenile dementia, Leukemia, Migraine disorders, Mental disorders	Binge eating disorder, Diabetic neuropathies, Marijuana abuse, Neuropathy, Crohn disease
Asian	Heart diseases, Food poisoning, Obesity, Herpes NOS, Stress	Kidney diseases, Myopia, Fatigue, Hemorrhage, Hypersensitivity
Hispanic	Headache, Fibromyalgia, Sleeplessness, Mental depression, Insomnia adverse event	Herpes zoster disease, Diarrhea, Suicide, Lupus vulgaris, Osteoporosis

Location

Table 9 summarizes the top disorder results for all sources. Focusing on drugs and forums, which have been shown to have more useful information regarding one's health [13], our key finding is that users in the Northeast talk more about traditional physical disorders including diabetes and heart conditions, users in the Midwest discuss about weight loss, users in the South about fibromyalgia, and users in the West discuss mental disorders and addictive behaviors.

The other location-based results including sentiment, emotions, top distinctive terms, and top distinctive drugs exhibit less variance among the location groups, and hence, we report them in Tables C.9, C.10, C.11, and C.12 of Multimedia Appendix 1, as the variations across locations are not significant.

Writing Level

Table 10 summarizes the emotion results for all sources. For shortness, only 3 emotions are listed here: anger, trust, and anticipation, as the other 3 (fear, disgust, and surprise), are complementary to these, respectively. We see that users with lower writing level express more anger, with the exception of drug review websites, whereas people with higher writing level express less anger. Due to low variance among writing levels, the other results for writing level including sentiment, top distinctive terms, top distinctive disorders, and top distinctive drugs can be found in Tables C.13, C.14, C.15, and C.16 of Multimedia Appendix 1, respectively.

XSL•FO

RenderX

Table 9. Top 5 distinctive disorders by location.

Location	TwitterHealth	Google+Health	Drugs	Forums
Northeast	N/A	Inflammatory bowel diseases, Crohn disease, Occupant of van injured in transport accident, Kidney diseases, Prostate carcinoma	Asleep, Seizures, Patient outcome—died, Memory observations, Fatigue	Diabetes, Atrial fibrillation, Gastroesophageal reflux disease, ACHE, Lupus vulgaris
Midwest	Migraine disorders, Primary malignant neoplasm	Confusion, Marijuana abuse, Van der Woude syndrome, Injury wounds, Cataract	Hemorrhage, Body weight decreased, Hot flushes, Weight loss adverse event, Xerostomia	Asthma, Migraine disorders, Pressure (finding), Autistic disorder, Cerebrovascular accident
South	N/A	Diabetic neuropathies, Binge eating disorder, Neuropathy, Alzheimer's disease pathway KEGG, Diarrhea	Fibromyalgia, Drowsiness, Edema, Pruritus, Manic	Codependency, Shot (injury)
West	Presenile dementia, Heart diseases, Mental suffering, Herpes NOS, Obesity	Sexually transmitted diseases, Bipolar disorder, Myocardial infarction, Vitality, Acquired immunodeficiency syndrome	Post-traumatic stress disorder, Sleeplessness, Anxiety attack, Addictive behavior, Suicidal	Marijuana abuse, Addictive behavior, codependency, Autistic disorder, Lupus erythematosus, systemic

Table 10. Emotion for each demographic grouped by source.

Writing level	TwitterHealth			Google+Health			Drugs			Forums		
	Anger (%)	Trust (%)	Anticipation (%)	Anger (%)	Trust (%)	Anticipation (%)	Anger (%)	Trust (%)	Anticipation (%)	Anger (%)	Trust (%)	Anticipation (%)
0-5	N/A	N/A	N/A	38.2[a]	68.5	71.2[a]	31.6[a]	66.8[a]	72.9[a]	34.3[a]	78.1[a]	71.5[a]
6-9	41.0[a]	44.3[a]	75.3[a]	34.6[a]	67.9	75.8[a]	31.4[a]	67.7[a]	73.2[a]	31.7[a]	77.8[a]	72.6[a]
10-16	34.1[a]	55.2[a]	81.9[a]	31.0[a]	66.6	79.1[a]	29.9[a]	73.1[a]	72.9[a]	27.4[a]	77.2[a]	75.4[a]

[a]Represents the values with high significance ($P \leq .05$) compared with the union of the other age groups.

Discussion

Notable Results

Our results provide valuable information that can help reach the right demographic group for each health condition. For example, to reach young users (aged 0-17 years) with ADHD, one should go to drug review websites. This finding can be a result of the increased percentage of children with ADHD recently (9%), compared with 2000 when it was 7% [44]. Similarly, to reach users of age group 18-34 with sleep disorder, one should go to Google+. We also found that the age group of 35-44 years discusses drugs associated with MS disorder, which agrees with the average age of clinical onset of MS, which is 30-33 years, and the average age of diagnosis, which is 37 years [45]. Because older age groups as our results show tend to discuss chronic diseases such as diabetes, heart problems, and cholesterol, health professionals and educators can target these groups in drug review websites and health Web forums to increase national awareness and decrease disease-related deaths.

Furthermore, a surprising finding is that, despite the fact that women suffer from back pain more than men [46], our results show that men discuss back pain more than women. Because 76% of all adults who have HIV are men [47], our results support this fact where HIV is one of the top discussed topics in TwitterHealth.

Our results also found that users in Western states discuss mental disorders and addictive behaviors including alcohol and marijuana as Table C.11 of Multimedia Appendix 1 shows. This finding is associated with the fact that 5 of top 10 states with high marijuana use are in the West area [48]. Midwest users discuss weight loss more than the other regions according to our results, which can be related to the fact that the Midwest is the second (slightly trailing the South) highest region in terms of obesity, with more than 25% of the adults being obese (body max index of 30+) [49].

Applications

There are several ways to leverage our results. Our findings can help health care providers and public health officials create targeted and effective educational campaigns, guide advertisers for different topics discussed by different demographic groups, help funding agencies allocate their research funds to have a larger impact on the society's top health issues, and help understand health disparities in Web-based health social media.

For instance, to reach pregnant or trying-to-get-pregnant women, advertisers should go to health Web forums and drug review websites instead of Google+ for advertising related products. This finding is supported by the fact that drug review websites and health Web forums are dominated by female users [21]. Also, this finding may indicate that there is a need for more definitive and authoritative sources of such information.

XSL·FO
RenderX

Our results can also help understand health disparities in Web-based health social media. Users with higher writing level are less angry when discussing health-related issues, which may be linked to the fact that people with lower level of education receive lower quality of health care [50] and have higher mortality rate [14].

These demographics-specific findings can be used in targeted educational campaigns, which are recently becoming the focus of several research efforts. As an example, Whittaker [51] shows how a smoking cessation intervention using mobile phones for young adults can be effective by sending general health videos messages and setting a quit date. Furthermore, Opel et al [52] show how social marketing can be used to increase immunization rates, where they explained how social marketing techniques can capture attention and motivate the targeted population to change. Patel et al [53] performed a systemic review to evaluate the effect of applications of contemporary social media on clinical outcomes in chronic disease. The study shows that providing social, emotional, and experiential support in current social media can help improve the patient care. Valle et al [54] evaluated a Facebook-based intervention that aims to increase the physical activity of young adult cancer survivors, which shows a potential for increasing the physical activity compared with Facebook-based self-help. A review of health interventions in Web-based social networks is presented in the study by Maher et al [55] where it is shown that several studies included in the systematic review reported significant improvement in health behavior or outcomes.

Limitations

For the general social networks, Google+ and Twitter, we used 276 health keywords and phrases as we described in the Methods section to filter the posts. These keywords and phrases miss some consumer phrases or abbreviations, such as ivf (in vitro fertilization) and iui (Intrauterine insemination). Unfortunately, we must select a relatively small set of keywords, given the rate constraints of the APIs of the social media.

Owing to the fact that ethnicity was estimated using a classifier [21], we were not able to confidently compute the ethnicity of enough users to have reliable results for several cases. For that, we omit results for black users. Furthermore, we do not report ethnicity results for drug review websites and health Web forums because these sources do not provide users' last names. Another limitation is self-selection bias because all demographic attributes (explicitly reported or classified) are reported by users. For instance, a user may choose to report age or last name

(which is used to classify ethnicity). For example, people who trust the opinion of other users or experts participate more in social networks, whereas people who have less trust might not share their private experiences.

For extracting medical concepts, we do not handle all abbreviations. We handle some cases through manual rules, for example, Metamap would map "I" to iodine. Also, MetaMap is not perfect for annotating social media posts; thus, we removed annotations that look incorrect as the previous example. Moreover, when computing the top distinctive terms, we do not handle variations of terms, that is, "iui" and "Intrauterine insemination" are considered different terms. We do a manual postprocessing to address this issue for the top results. In measuring the sentiment of posts, the sentiment lexicon "SentiWordNet" was not built specifically for social or medical text. For example, some words such as "omg" or "lol" are not mapped to any word in the lexicon; thus, not all terms in the posts are assigned a sentiment.

Conclusion

We analyzed the content of Web-based health social media based on users' demographics. Three different types of Web-based health social media were considered: social networks, drug review websites, and health Web forums. For each demographic attribute—gender, age, ethnicity, location, and writing level—we evaluated sentiment and emotion, and we extracted top distinctive terms and medical concepts, specifically disorders and drugs. Our results are both expected and surprising and show several key findings for each demographic attribute. For example, the dominant topic for female users in drug review websites and health Web forums is pregnancy, whereas for male users, it is cardiac problems, HIV, and back pain. Attention-deficit hyperactivity disorder and depression-related drugs are the main topics discussed by younger users (0-17 years), MS drugs are discussed more by users of age 35-44 years, and alcohol and smoking are mainly discussed by middle-aged users (45-64 years). Users from the Northeast United States talk about physical disorders, whereas users from the West United States talk about mental disorders and addictive behaviors. Finally, users with higher writing level express less anger in their posts. These key findings can help experts reach the right users in many ways, including creating targeted and effective educational campaigns by health care providers, advertising related products, allocating funds for the right research by funding agencies, and understanding health disparities in Web-based health social media.

Authors' Contributions

All authors contributed substantially to this work. They designed and performed the analysis and approved the final version of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Web-Based Social Network Summary, Health-Related Keywords, Results, and Statistical Significance Tests.

[PDF File (Adobe PDF File), 191KB - jmir_v18i6e148_app1.pdf]

References

1. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013 Apr;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]
2. Fox S, Jones S. The Social Life of Health Information. 2009 Jun 11. Pew Research Center URL: http://www.pewinternet.org/2009/06/11/the-social-life-of-health-information [accessed 2016-05-19] [WebCite Cache ID 6cpiXv3kp]
3. Kane GC, Fichman RG, Gallaugher J, Glaser J. Community relations 2.0. Harv Bus Rev 2009 Nov;87(11):45-50. [Medline: 19891388]
4. Chou WS, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. J Med Internet Res 2009;11(4):e48 [FREE Full text] [doi: 10.2196/jmir.1249] [Medline: 19945947]
5. Kontos EZ, Emmons KM, Puleo E, Viswanath K. Communication inequalities and public health implications of adult social networking site use in the United States. J Health Commun 2010;15 Suppl 3:216-235 [FREE Full text] [doi: 10.1080/10810730.2010.522689] [Medline: 21154095]
6. Unified Medical Language System (UMLS). URL: https://www.nlm.nih.gov/research/umls/ [accessed 2015-11-06] [WebCite Cache ID 6cpj8xvk5]
7. Luo JS. Social Networking, Health 2.0, and Beyond. In: Smith BN, editor. Information Technology Essentials for Behavioral Health Clinicians (Health Informatics). London: Springer; 2011:119-131.
8. Hackworth BA, Kunz MB. Health care and social media: building relationships via social networks. Academy of Health Care Management Journal 2011;7(2):1-14 [FREE Full text]
9. Grajales FJ, Sheps S, Ho K, Novak-Lauscher H, Eysenbach G. Social media: a review and tutorial of applications in medicine and health care. J Med Internet Res 2014;16(2):e13 [FREE Full text] [doi: 10.2196/jmir.2912] [Medline: 24518354]
10. ECRI. 2011 Nov. Social media in healthcare URL: http://kcmsociety.org/wp-content/uploads/2013/04/ECRI-social-media.pdf [accessed 2015-02-01] [WebCite Cache ID 6W1rk2Yt2]
11. Denecke K, Nejdl W. How valuable is medical social media data? Content analysis of the medical web. Information Sciences 2009 May 30;179(12):1870-1880. [doi: 10.1016/j.ins.2009.01.025]
12. Lu Y, Zhang P, Liu J, Li J, Deng S. Health-related hot topic detection in online communities using text clustering. PLoS One 2013 Feb;8(2):e56221 [FREE Full text] [doi: 10.1371/journal.pone.0056221] [Medline: 23457530]
13. Wiley MT, Jin C, Hristidis V, Esterling KM. Pharmaceutical drugs chatter on Online Social Networks. J Biomed Inform 2014 Jun;49:245-254. [doi: 10.1016/j.jbi.2014.03.006] [Medline: 24637141]
14. Krueger PM, Tran MK, Hummer RA, Chang VW. Mortality Attributable to Low Levels of Education in the United States. PLoS One 2015 Jul;10(7):e0131809 [FREE Full text] [doi: 10.1371/journal.pone.0131809] [Medline: 26153885]
15. Pew Research Center. 2013 Dec 27. Social networking fact sheet URL: http://www.pewinternet.org/fact-sheets/social-networking-fact-sheet/ [accessed 2016-05-26] [WebCite Cache ID 6W22L3O2w]
16. eMarketer. 2013 Apr 25. Hispanic shoppers bring social, mobile habits to the grocery aisle URL: http://www.emarketer.com/Article/Hispanic-Shoppers-Bring-Social-Mobile-Habits-Grocery-Aisle/1009839 [accessed 2016-05-26] [WebCite Cache ID 6W21kTHSr]
17. Mislove A, Lehmann S, Ahn YY, Onnela JP, Rosenquist JN. Understanding the demographics of Twitter users. ICWSM 2011;11:5 [FREE Full text]
18. Mandel B, Culotta A, Boulahanis J, Stark D, Lewis B, Rodrigue J. A demographic analysis of online sentiment during hurricane irene. In: Proceedings of the Second Workshop on Language in Social Media. 2012 Presented at: Association for Computational Linguistics; 2012; Stroudsburg, PA, USA p. 27-36.
19. Liu Y, Xu S, Yoon H, Tourassi G. Extracting patient demographics and personal medical information from online health forums. AMIA Annu Symp Proc 2014;2014:1825-1834 [FREE Full text] [Medline: 25954455]
20. Anderson-Bill ES, Winett RA, Wojcik JR. Social cognitive determinants of nutrition and physical activity among web-health users enrolling in an online intervention: the influence of social support, self-efficacy, outcome expectations, and self-regulation. Journal of Medical Internet Research 2011 Mar 17;13(1):e28 [FREE Full text] [doi: 10.2196/jmir.1551] [Medline: 21441100]
21. Sadah SA, Shahbazi M, Wiley MT, Hristidis V. A Study of the Demographics of Web-Based Health-Related Social Media Users. Journal of Medical Internet Research 2015;17(8):e194 [FREE Full text] [doi: 10.2196/jmir.4308] [Medline: 26250986]

XSL•FO

RenderX

22.   RxList - The Internet Drug Index for prescription drugs, medications and pill. URL: http://www.rxlist.com/script/main/hp.asp [accessed 2015-02-02] [WebCite Cache ID 6W24Layfl]

23.   Twitter. URL: https://twitter.com/ [accessed 2016-05-19] [WebCite Cache ID 6hd6ik8QF]

24.   Google+. URL: https://accounts.google.com/ServiceLogin?service=oz&passive=1209600&continue=https://plus.google.com/?gpsrc%3Dgplp0 [accessed 2016-05-19] [WebCite Cache ID 6hd6ooiY6]

25.   Drugs.com | Prescription Drug Information, Interactions & Side Effects. URL: http://www.drugs.com/ [WebCite Cache ID 6W23lHwlt]

26.   Treatments: reviews of drugs, therapies and remedies by everyday people - DailyStrength. URL: http://www.dailystrength.org/treatments [accessed 2015-02-02] [WebCite Cache ID 6W23nXoJa]

27.   WebMD. WebMD Drugs & Treatments - Medical Information and user ratings on prescription drugs and over-the-counter (OTC) medications URL: http://www.webmd.com/drugs/index-drugs.aspx [accessed 2015-02-02] [WebCite Cache ID 6W23rKUfB]

28.   Drugs. Medical Questions Answered - Drugs URL: http://www.drugs.com/answers/ [accessed 2015-02-02] [WebCite Cache ID 6W23vDAi5]

29.   DailyStrength. Online Support Groups - DailyStrength URL: http://www.dailystrength.org/support-groups [accessed 2015-02-02] [WebCite Cache ID 6W23ztpRe]

30.   WebMD. Better information. Better health URL: http://www.webmd.com/ [WebCite Cache ID 6W247u0HQ]

31.   Twitter. The Streaming APIs | Twitter Developers URL: https://dev.twitter.com/streaming/overview [accessed 2015-06-15] [WebCite Cache ID 6ZHpD7gjO]

32.   Developers.google. Google+ API | Google+ Platform for Web | Google Developers URL: https://developers.google.com/+/api/ [accessed 2015-02-02] [WebCite Cache ID 6W24x2YYc]

33.   jsoup. jsoup Java HTML Parser, with best of DOM, CSS, and jquery URL: http://jsoup.org/ [accessed 2015-02-02] [WebCite Cache ID 6W2528eFu]

34.   Readabilityformulas. The flesch grade level readability formula URL: http://www.readabilityformulas.com/flesch-grade-level-readability-formula.php [accessed 2015-02-02] [WebCite Cache ID 6W22mEB3h]

35.   Baccianella S, Esuli A, Sebastiani F. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. 2010 May 17 Presented at: LREC; May 17, 2010; Malta p. 2200-2204.

36.   Mohammad SM, Turney PD. Crowdsourcing a word–emotion association lexicon. Computational Intelligence, 2013 Aug 1;29(3):436-465. [doi: 10.1111/j.1467-8640.2012.00460.x]

37.   Mohammad SM, Yang TW. Tracking Sentiment in Mail: How Genders Differ on Emotional Axes. In: Proceedings of the 2nd workshop on computational approaches to subjectivity and sentiment analysis. 2011 Presented at: Association for Computational Linguistics; June 19, 2011; Portland, Oregon, USA p. 70-79 URL: http://aclweb.org/anthology-new/W/W11/W11-1709.pdf

38.   Ohana OB, Tierney B. Sentiment classification of reviews using SentiWordNet. In: IT & T Conference. 2009 Oct 01 Presented at: Dublin Institute of Technology; October 22, 2009; Dublin, Ireland p. 2009.

39.   Toutanova K, Klein D, Manning C. Feature-rich part-of-speech tagging with a cyclic dependency network. In: Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology. 2003 Presented at: Association for Computational Linguistics; 2003; Stroudsburg, PA, USA p. 173-180. [doi: 10.3115/1073445.1073478]

40.   Tartarus. Porter Stemming Algorithm URL: http://tartarus.org/martin/PorterStemmer/ [WebCite Cache ID 6cpl5KN6N]

41.   Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. Nucleic Acids Res 2004 Jan 1;32(Database issue):D267-D270 [FREE Full text] [doi: 10.1093/nar/gkh061] [Medline: 14681409]

42.   Aronson AR. Effective mapping of biomedical text to the UMLS Metathesaurus: the MetaMap program. Proc AMIA Symp 2001:17-21 [FREE Full text] [Medline: 11825149]

43.   Denecke K, Soltani N. The Burgeoning of Medical Social-Media Postings and the Need for Improved Natural Language Mapping Tools. In: Where Humans Meet Machines: Innovative Solutions for Knotty Natural-Language Problems. New York: Springer; 2013:27-43.

44.   Akinbami LJ, Liu X, Pastor PN, Reuben CA. Attention deficit hyperactivity disorder among children aged 5-17 years in the United States, 1998-2009. NCHS Data Brief 2011 Aug(70):1-8 [FREE Full text] [Medline: 22142479]

45.   Statistic Brain. Multiple Sclerosis Statistics URL: http://www.statisticbrain.com/multiple-sclerosis-statistics/ [accessed 2015-11-06] [WebCite Cache ID 6cplolLte]

46.   Statista. Percentage of adults in the U.S. with low back pain from 1997 to 2013, by gender URL: http://www.statista.com/statistics/188858/adults-in-the-us-with-low-back-pain-by-gender-since-1997/ [accessed 2015-11-06] [WebCite Cache ID 6cplm2Pox]

47.   CDC. HIV Among Men in the United States URL: http://www.cdc.gov/hiv/group/gender/men/index.html [accessed 2016-05-20] [WebCite Cache ID 6fzj6kzOK]

48.   Rawes E. Money & Career CheatSheet. 2015 May 1. 10 States with the Highest Marijuana Use URL: http://www.cheatsheet.com/personal-finance/10-states-with-the-highest-marijuana-use.html/ [accessed 2015-11-06] [WebCite Cache ID 6cplxkbG7]

49.   The State of Obesity. Adult Obesity in the United States URL: http://stateofobesity.org/adult-obesity/ [accessed 2015-11-06] [WebCite Cache ID 6cpm04NQD]

50.   Smith JC, Medalia C. Health Insurance Coverage in the United States: 2014. US Census Bureau 2015 Sep:60-253.

51.   Whittaker R. Smoking cessation intervention for young adults using multimedia mobile phones: development and effectiveness. Development and effectiveness (Doctoral dissertation, ResearchSpace@ Auckland) 2011:1-366 [FREE Full text]

52.   Opel DJ, Diekema DS, Lee NR, Marcuse EK. Social marketing as a strategy to increase immunization rates. Arch Pediatr Adolesc Med 2009 May;163(5):432-437. [doi: 10.1001/archpediatrics.2009.42] [Medline: 19414689]

53.   Patel R, Chang T, Greysen SR, Chopra V. Social Media Use in Chronic Disease: A Systematic Review and Novel Taxonomy. Am J Med 2015 Dec;128(12):1335-1350. [doi: 10.1016/j.amjmed.2015.06.015] [Medline: 26159633]

54.   Valle CG, Tate DF, Mayer DK, Allicock M, Cai J. A randomized trial of a Facebook-based physical activity intervention for young adult cancer survivors. J Cancer Surviv 2013 Sep;7(3):355-368 [FREE Full text] [doi: 10.1007/s11764-013-0279-5] [Medline: 23532799]

55.   Maher CA, Lewis LK, Ferrar K, Marshall S, De BI, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]

Abbreviations

AA: alcoholic anonymous
API: Application Program Interface
ADHD: attention-deficit hyperactivity disorder
BC: birth control
BP: blood pressure
BFN: big fat negative
BFP: big fat positive
COPD: chronic obstructive pulmonary disease
IUI: intrauterine insemination
IVF: in vitro fertilization
MS: multiple sclerosis
NLP: Natural Language Processing
PCOS: polycystic ovary syndrome
PD: panic disorder
RA: rheumatoid arthritis
Stbx: Soon-to-be-ex
TTC: trying to conceive
UMLS: Unified Medical Language System

XSL•FO
RenderX

Original Paper

Beyond Traditional Newspaper Advertisement: Leveraging Facebook-Targeted Advertisement to Recruit Long-Term Smokers for Research

Lisa Carter-Harris[1*], PhD, RN, ANP-C; Rebecca Bartlett Ellis[1*], PhD, RN, CNS; Adam Warrick[2]; Susan Rawl[1*], PhD, RN, FAAN

[1]Indiana University, School of Nursing, Indianapolis, IN, United States
[2]Indiana University Bloomington, Bloomington, IN, United States
[*]these authors contributed equally

Corresponding Author:
Lisa Carter-Harris, PhD, RN, ANP-C
Indiana University
School of Nursing
600 Barnhill Drive
W427
Indianapolis, IN, 46202
United States
Phone: 1 317 274 2043
Fax: 1 317 274 2411
Email: lcharris@iu.edu

Abstract

Background: Smokers are a stigmatized population, but an important population to reach for the purpose of research. Therefore, innovative recruitment methods are needed that are both cost-effective and efficacious in recruiting this population.

Objective: The aim of the present article was to evaluate the feasibility of Facebook-targeted advertisement to recruit long-term smokers eligible for lung cancer screening for a descriptive, cross-sectional survey.

Methods: A social media recruitment campaign was launched using Facebook-targeted advertisement to target age and keywords related to tobacco smoking in the Facebook users profile, interests, and likes. A 3-day newspaper advertisement recruitment campaign was used as a comparison. The study that used both recruitment methods aimed to test the psychometric properties of 4 newly developed lung cancer screening health belief scales. Data were collected via cross-sectional survey methodology using an Web-based survey platform.

Results: The Facebook-targeted advertisements were viewed 56,621 times over an 18-day campaign in 2015 in the United States. The advertisement campaign yielded 1121 unique clicks to the Web-based survey platform at a cost of \$1.51 per completed survey. Of those who clicked through to the study survey platform, 423 (37.7%) consented to participate; 92 (8.2%) dropped out during completion of the survey yielding a final study pool of 331 completed surveys. Recruitment by newspaper advertisement yielded a total of 30 participants in response to a 3-day advertisement campaign; recruitment efficacy resulted in 10 participants/day at \$40.80 per completed survey. Participants represented current (n=182; 51%) and former smokers (n=178; 49%) with a mean age of 63.4 years (SD 6.0). Cost of the advertisement campaign was \$500 total for the 18-day campaign.

Conclusions: Recruitment by Facebook was more efficacious and cost-effective compared with newspaper advertisement. Facebook offers a new venue for recruitment into research studies that offer the potential for wider reach at a lower cost while providing privacy and flexibility for potential study participants. The study's findings extend recent work of other researchers who have demonstrated Facebook's utility with younger smokers, and Facebook is an effective tool to recruit older smokers. Furthermore, Facebook is a cost-effective alternative to traditional newspaper advertisement offering a new, affordable venue to recruit large numbers of older smokers efficiently.

XSL•FO
RenderX

Introduction

Lung cancer is the deadliest cancer-related diagnosis worldwide regardless of gender or ethnicity with most patients diagnosed at an advanced stage [1]. For the first time, there is a screening test for high-risk individuals (*defined as current or former smokers who have quit within the past 15 years who have a 30 pack-year tobacco smoking history*) [2,3]. Lung cancer screening with low-dose computed tomography (LDCT) in long-term smokers has been shown to decrease relative lung cancer-related mortality by 20% [2]. An LDCT is a newer form of a computed tomography scan that uses lower doses of radiation to take a series of 3D radiographs of the lungs. These images are detailed and can show early-stage lung cancers that may be too small for conventional chest radiographs to detect [2]. In response to empirical findings from the National Lung Screening Trial, the US Preventive Services Task Force issued lung cancer screening guidelines recommending annual LDCT of the chest for high-risk individuals in 2013 [4].

As lung cancer screening becomes more widely implemented, participation is likely to be influenced by many factors, including individual-, provider-, and health care system-related variables. To determine factors that influence lung cancer screening participation, understanding perspectives of individuals eligible for screening is essential. Current lung cancer screening guidelines target long-term current and former smokers [3]. However, recruitment of smokers can be a challenge, and such research can be limited by the ability to access this target population. Smoking-related stigma has been implicated in timing of medical help-seeking behavior in symptomatic individuals later diagnosed with lung cancer [5] and quality of life in current and former smokers diagnosed with lung cancer [6]. Smoking-related stigma may serve as a barrier to recruitment of this important population. Current and former smokers may worry about being blamed for their smoking history as well as feel like social outcasts for smoking, increasing a sense of internalized stigma [7,8]. They may also fear having to endure a lecture from their health care provider about their current smoking status. For researchers wishing to recruit smokers, traditional methods such as face-to-face recruitment and fliers placed in high-traffic areas may not be as successful as recruitment targeting other demographics for research studies [9]. In addition, newspaper advertisement may be costly. Facebook is a relatively new venue for recruitment into research studies and offers the potential for wider reach at a lower cost while providing privacy and flexibility for potential study participants. Innovative methods are needed that are both cost-effective and efficient in recruiting this potentially hard-to-recruit population.

Facebook has previously been established as a viable option to recruit young adults into health-related research [10-17] and may be a successful recruitment tool for older adults. Facebook has been a successful recruitment tool to reach adolescents and young adults for a range of study purposes including exploring mental health issues [13,18], examining pubescent hormonal effects in early adolescence [19], and recruiting for a variety of Web-based intervention studies [18,20,21]. Facebook may also be a beneficial resource for retention in longitudinal studies as previous studies have demonstrated its utility in retaining adolescents via social media [14-15]. Facebook has also been a successful recruitment strategy for reaching "hard-to-reach" populations such as men who have sex with men [22] and exploring taboo topics with adolescents and young adults such as abortion [16]. Finally, studies targeting young adult smokers have demonstrated Facebook advertisements as a successful recruitment tool [9,17,23,24]. Therefore, our study sought to determine whether Facebook would be a successful recruitment tool in a new domain: older, long-term smokers.

As of March 2015, Facebook reported 936 million active daily users worldwide [25]. In the United States, 71% of adults use Facebook [26]. Facebook has also grown in use among older individuals and is visited by 63% of individuals aged 50 to 64 years and by 56% of individuals aged 65 years and older in a typical day [26]. Facebook offers the ability for the researcher to market a recruitment campaign of advertisements targeted by age, location, and keywords identified in the profile or interest list of potential participants. This strategy of targeted advertisement has the potential to engage current and former smokers as research participants while maintaining a sense of privacy for the individual potentially decreasing the perception of associated stigma. The purpose of this article was to describe the method by which a national sample of older long-term current and former smokers was successfully recruited into a descriptive survey study using Facebook-targeted advertisement.

Methods

Sample

This Facebook sample was recruited as part of a larger overall study to psychometrically test 4 newly developed scales to measure health beliefs about lung cancer screening. The findings of the psychometric study are published [27]. For the larger study, we aimed to recruit men and women who were eligible for lung cancer screening and included individuals between the ages of 55 and 77 years who were current or former smokers who had quit within the past 15 years and had a 30 pack-year tobacco smoking history. Pack-year is defined as the number of packs of cigarettes smoked per day multiplied by the number of years smoked. Individuals diagnosed with lung cancer were excluded from the study. It should be noted that the age range for the inclusion criteria of the psychometric study was 55 to 77 years. However, Facebook-targeted advertisement does not currently offer the ability to narrow age range in the 65-years-and-older category. Therefore, the age targeted for the purposes of the Facebook advertisement was 55 years and older with the ability to analyze Facebook advertisement metrics by 55 to 64 years and 65 years and older as discrete ranges. Two recruitment strategies were used: (1) a Facebook advertisement campaign and (2) a newspaper advertisement. The required sample size, based on the planned statistical analyses to evaluate

XSL•FO

RenderX

the psychometric properties of the lung cancer screening health belief scales, was 300 completed surveys. Institutional review board approval was obtained from Indiana University before recruitment of study participants.

Procedures

Facebook Advertisement Campaign

A social media recruitment campaign was launched using Facebook-targeted advertisement over an 18-day period. Facebook uses 2 types of advertisements that can appear in different locations on the screen depending on platform used to access Facebook (ie, desktop or mobile app). To best understand these advertisement locations, key terminology specific to Facebook advertisement must be explained to include: (1) newsfeed; (2) impressions; (3) reach; and (4) clicks to website. A *newsfeed* is a constantly updating list of stories in the middle of the Facebook user's homepage. The newsfeed can include status updates, photos, videos, links, application activity, and likes from people, pages, and groups followed by the Facebook user. Advertisements also appear in a Facebook user's newsfeed. The 2 types of advertisements specific to the desktop platform include: (1) standard desktop advertisements that appear in the right hand column next to the newsfeed on the Facebook user's homepage and (2) newsfeed advertisements that appear in the constantly updating newsfeed located in the middle of the Facebook user's homepage. For mobile app users, the newsfeed advertisement appears only in the middle of the constantly updating newsfeed. An *impression* refers to the number of times the advertisement entered the screen for the first time (ie, is served to someone) either in their desktop newsfeed, mobile newsfeed, or as a right hand column advertisement. *Reach* refers to the number of people to whom the advertisement was shown. A *click to website* refers to a unique Facebook user clicking the weblink embedded in the Facebook advertisement that is redirected to the advertised website. Another method to reach potential study participants in Facebook is to target the Facebook advertisement to the audience network of a Facebook group page. An *audience network* refers to individuals who have liked a Facebook group page. In the case of this study, the researchers set up a Facebook

group page called Healthy Lungs Initiative with the purpose of providing general lung health information to recruit individuals interested in lung health issues. The Healthy Lungs Initiative Facebook group page was set up at the start of the study and remains active.

A recruitment advertisement targeting Facebook users whose age were 55 years and older and lived in the United States was used. Keywords used in targeting the selected group for our recruitment included "tobacco," "tobacco smoking," "smoking," "smoking cessation," "cigarettes," and "electronic cigarettes" using both standard desktop and newsfeed advertisements (mobile and desktop). When we defined our audience during creation of the advertisement using these keywords, age, and location, we had a potential reach of 910,000 people. Please see Figure 1 for metrics related to potential reach specific to each stage of the keyword targeting process. Standard desktop and newsfeed advertisements included a short headline, a picture of an individual getting an LDCT scan for lung cancer screening, a short description of the study, and a link to the study's Web-based survey (see Figure 2). The study's Web-based survey was conducted through an external website using the Research Electronic Data Capture (REDCap) software system. REDCap is a secure Web-based application for building and managing Web-based surveys and databases. As previously mentioned, the advertisement was connected to a Facebook group page set up specifically for this study by the researchers called Healthy Lungs Initiative. The advertisement was reviewed and approved by Facebook staff before being released per Facebook policy. Dates for the recruitment campaign were specified, and a lifetime spending limit for the recruitment campaign was set at $500 during the creation of the advertisement. In addition, we followed procedures to reduce participant misrepresentation as described by Kramer et al [28]. These include (1) prohibiting open access to the survey platform by embedding a weblink that redirects to a screening survey; (2) requiring screening questions to screen out individuals who do not meet the inclusion criteria; (3) incorporating a survey time stamp to examine initiation of survey versus survey completion time span; and (4) identifying item pairs that should be consistent.

Figure 1. Keyword Metrics of Potential Reach for Facebook Targeted Advertisement.
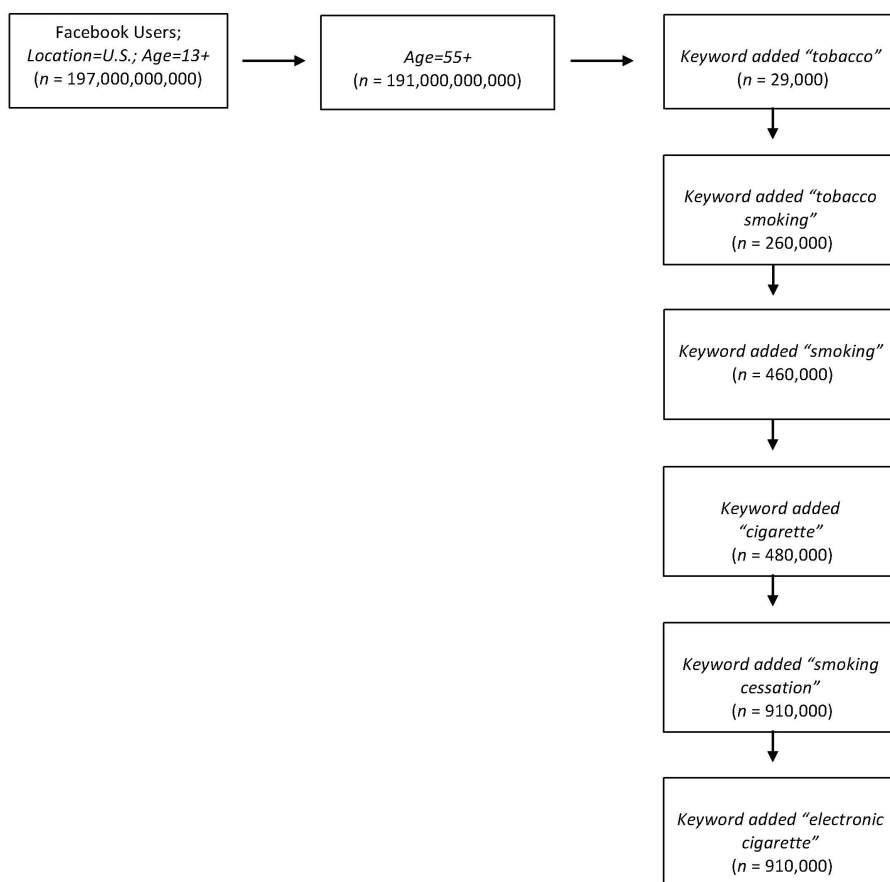
Figure 2. Facebook Recruitment Advertisement.

Newspaper Advertisement

A newspaper advertisement was placed in *The Indianapolis Star* newspaper in Indianapolis. Indiana is ranked 6th highest in adult smoking rates nationally with 21.9% of the state's adult men and women self-reporting as current smokers [1]. The advertisement was a 3.24" x 5" black and white announcement that ran for 3 consecutive days in the lifestyle section of the newspaper. The advertisement featured a black and white picture of diverse older adults of both genders and details related to eligibility criteria, what the study involved, and information to participate. The following 3 ways to participate were included in the advertisement: (1) visit a weblink to access the Web-based version of the survey; (2) email or call the research office to request a mailed copy of the survey; or (3) call the research office to complete the survey by telephone.

Results

The primary objective of this study was to evaluate the feasibility of Facebook-targeted advertisement to recruit long-term smokers eligible for lung cancer screening for a descriptive, cross-sectional survey. Although not a direct comparison in length of advertisement and geographic location, we present the newspaper advertisement results as a comparison of cost, time, and number of participants recruited per day.

Participant Characteristics

Of the 361 participants, the majority were recruited by Facebook advertisement (92%; n=331). Across both the recruitment methods, the majority were female (58%; n=211), non-Hispanic Caucasian (91%; n=327), high school graduates or higher (96%; n = 347), and equally representative of current (51%; n=182) and former smokers (49%; n=177). The mean age was 63.4 years (SD 6.0). Please see Table 1 for a complete list of participant sociodemographic characteristics in total and by recruitment method.

Table 1. Participant sociodemographic characteristics by recruitment method (N=361).

		Total (N=361)[b]	Newspaper[b](n=30)	Facebook[b](n=331)
Gender, n (%)				
	Female	211 (58)	17 (57)	194 (58)
	Male	150 (42)	13 (43)	137 (41)
Race or ethnicity, n (%)				
	Caucasian	327 (91)	22 (73)	305 (92)
	African-American	28 (8)	8 (27)	20 (6)
Education, n (%)				
	<High school	15 (4)	1 (3)	14 (4)
	High school graduate	104 (29)	6 (20)	98 (30)
	Some college	154 (43)	15 (50)	139 (42)
	College graduate	15 (4)	8 (27)	81 (24)
Annual income, n (%)				
	Less than $25,000	116 (32)	12 (40)	104 (31)
	$25,000-50,000	168 (47)	11 (37)	157 (47)
	More than $50,000	74 (21)	6 (20)	68 (21)
Insurance status, n (%)				
	Government sponsored	217 (60)	24 (80)	193 (58)
	Private insurance	132 (37)	5 (17)	127 (38)
	No insurance	13 (3)	1 (3)	12 (4)
Smoking status, n (%)				
	Current smoker	182 (51)	20 (67)	162 (49)
	Former smoker	177 (49)	10 (33)	167 (51)
Family history of lung cancer, n (%)				
	Yes	63 (18)	4 (13)	59 (18)
	No	299 (82)	26 (87)	273 (82)
Age (years), mean (SD) [a]		63.4 (6)	63.6 (6)	63.1 (6)

[a]SD: standard deviation.

[b]*Note.* Column percentages may not add to 100% due to missing data.

Newspaper Advertisement

One black and white newspaper advertisement ran over 3 consecutive days (Monday, Tuesday, and Wednesday) at a total cost of $1224. Subscribers aged 55 years and older (N=230,742) received the newspaper those 3 days. Therefore, the newspaper advertisement had a potential reach of 230,742 individuals aged 55 years and older who potentially saw the advertisement over the 3-day campaign. A total of 42 individuals responded to the advertisement; among them, 30 participants met the inclusion criteria, agreed to participate in the study, and completed the survey. Among the 30 participants, 21 (70%) completed the survey via the Web, 4 (13%) completed by telephone, and 5 (17%) completed a paper copy of the survey by email. Recruitment over the 3-day newspaper advertisement campaign averaged 10 participants per day (30 total participants or 3-day campaign=10), and cost of recruitment by newspaper advertisement was $40.80 per completed survey ($1224/30=$40.80).

Facebook Advertisement Respondents

During the 18-day Facebook recruitment campaign, 1121 unique Facebook users viewed, clicked on the advertisement, and were directed to the survey welcome page and screening survey. As depicted in the flow diagram in Figure 3 , of the 1121 people who viewed the survey welcome page and screening survey, 423 (37.7%) agreed to participate, but 92 (8.2%) people dropped out during the course of the survey. Although the survey was designed with the option to leave any item blank after proceeding past the screening survey, there were less than 5% missing data for all survey items. Recruitment by Facebook advertisement yielded 331 participants enrolled into the study.

Figure 3. Flow diagram of clicks to website, eligibility, consent to participate, and dropouts.
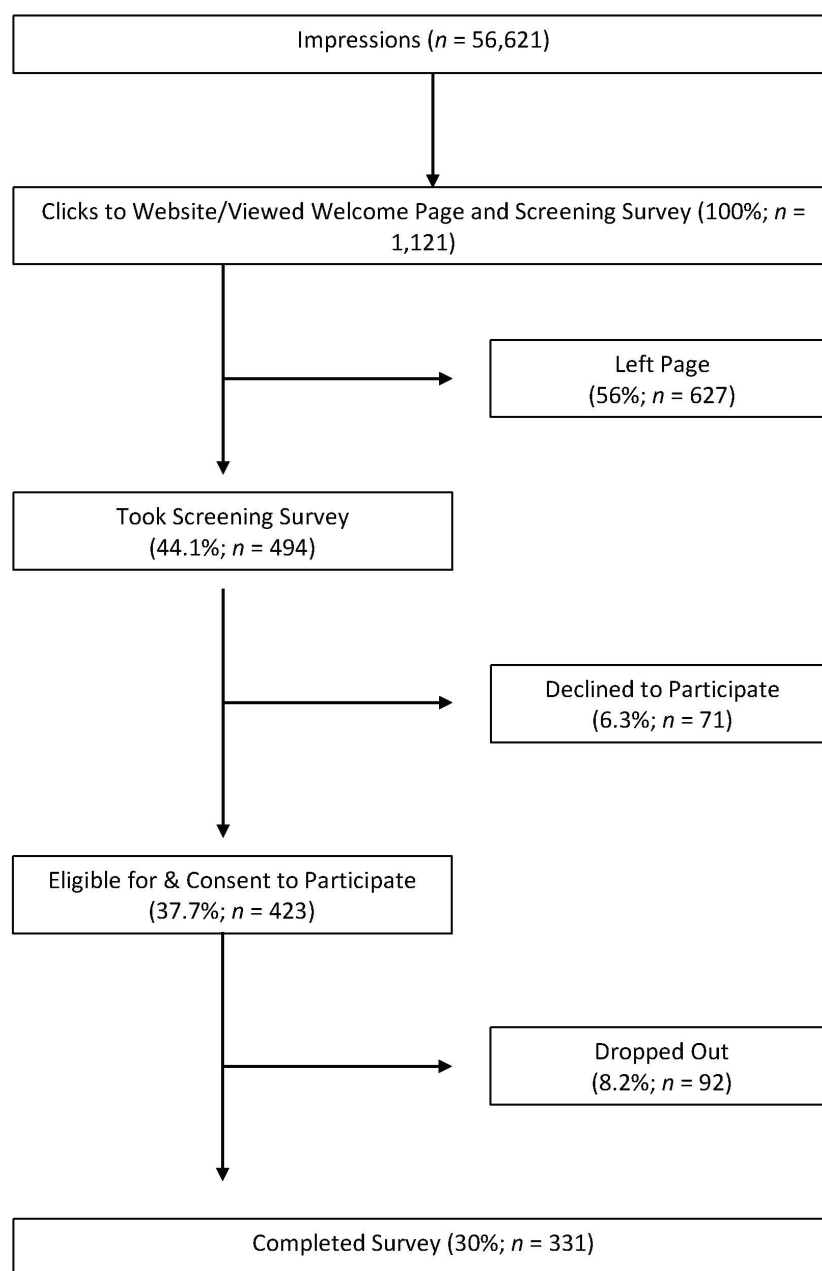
Please note. The denominator for the percentages is the number of individuals who clicked through to the website and viewed the welcome page and screening survey.

Facebook Recruitment Campaign Results

Over the 18-day Facebook recruitment campaign, the advertisement made 56,621 impressions yielding 1121 unique clicks to our Web-based survey at an overall cost of $500. Advertisements appeared on the right side of the desktop screen, within the desktop newsfeed as the potential participant scrolled, within the mobile newsfeed, or in the audience network of individuals who liked the Healthy Lungs Initiative Facebook group page. Newsfeed advertisements on a mobile device resulted in more unique users with 42,059 impressions and 894 unique clicks to the Web-based survey site than the other advertisement placements (see Table 2 for a complete list of

Facebook-advertising metrics). In addition, 7 unique clicks were generated from the Healthy Lungs Initiative Facebook group page. As mentioned previously, the Facebook group page was created, in conjunction with the Facebook-targeted advertisement for this study, as a platform for messages about general lung health issues. During the 18-day Facebook advertisement period, the Facebook group page obtained 29 likes. Recruitment over the 18-day Facebook advertisement campaign averaged 18 participants per day, and cost of recruitment was $1.51 per completed survey ($500/331=$1.51). See Figure 4 for number of participants recruited by Facebook advertisement per day.

Table 2. Characteristics of an 18-day Facebook-targeted advertisement campaign with Facebook-advertising metrics (N=331).

Variable		Reach	Unique Clicks to Website	Spent (cost per click)
Total Facebook users and budget		56,621	1121	$500
Advertisement placement				
	Desktop right side	4527	22	$0.32
	Desktop news-feed	7409	190	$0.30
	Mobile newsfeed	42,059	894	$0.48
	Audience network[a]	1015	7	$0.11
Gender				
	Women	35,581	722	$0.48
	Men	20,269	378	$0.40
	Unspecified	772	21	$0.29
Age, years				
	55-64	31,004	602	$0.44
	65+	25,617	519	$0.46

[a]Audience network refers to individuals who saw advertisement on Facebook page (Healthy Lungs Initiative).

Figure 4. Number of Participants Recruited from Facebook Each Day of the Study.
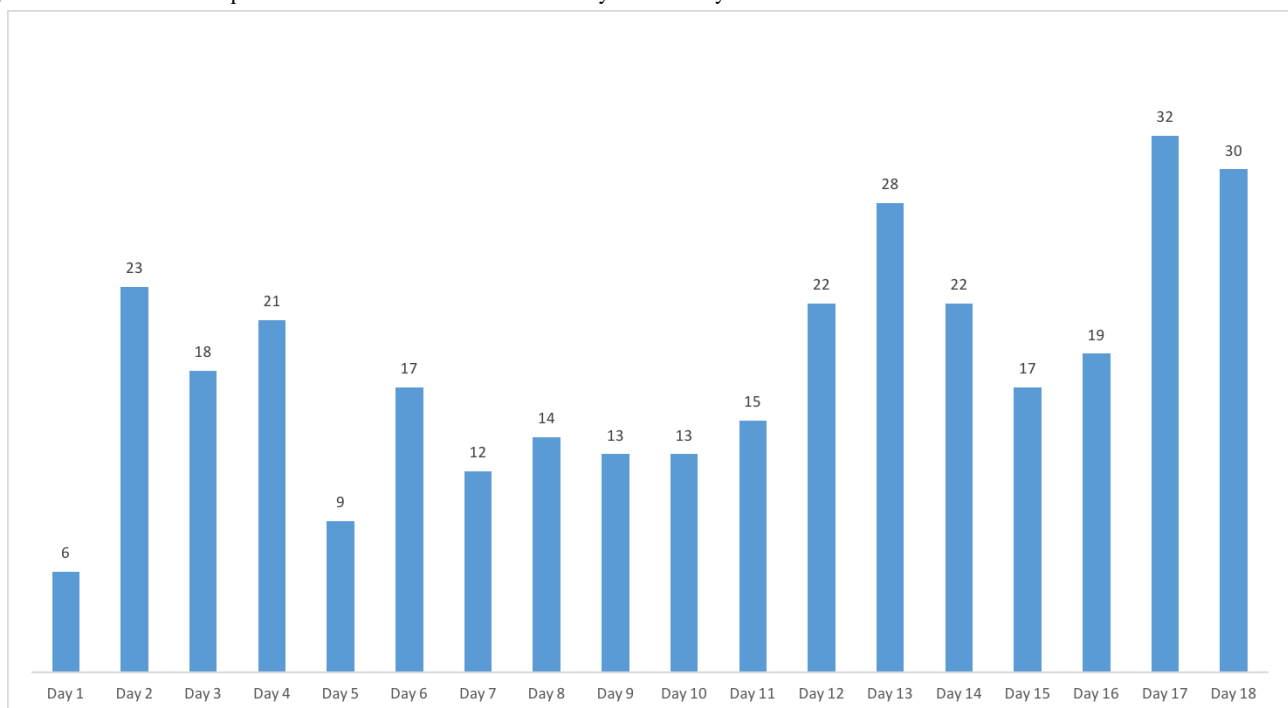
Discussion

Long-term current and former smokers can be challenging to reach and effectively recruit into research studies possibly secondary to perceived smoking-related stigma. However, this study demonstrated that Facebook is a viable recruitment method for this population. Facebook advertisement was more effective in cost and number of participants recruited compared with newspaper advertisement. Facebook advertisement cost was $1.51 per completed survey and averaged 18 recruited eligible participants per day compared with $40.80 per completed survey and an average of 10 recruited eligible participants per day with newspaper advertisement. In general, Facebook advertisement had minimal resource use, including time and cost related to research personnel. Management of the Facebook advertisement campaign involved less than 1 hour per day spent monitoring enrollment and stability of the embedded weblink to the survey. The newspaper advertisement, however, required greater personnel and time investment to answer telephone calls,

respond to voice mail messages, and administer the survey by telephone for participants who chose this option.

In addition to the low cost benefit, leveraging Facebook's ability to target specific study populations by demographic characteristics, location, and keywords in the Facebook users' profile, and interest list is of value to the researcher. This ability offers the opportunity to target recruitment advertisements based on multiple combinations of population characteristics important to their study. For this study, targeting by location, age, and keywords related to tobacco smoking and smoking cessation provided a potential reach of 910,000 unique Facebook users. We found this approach to be helpful in reaching our target audience of long-term current and former smokers because the keywords were likely personally relevant for the target population needed for the study. Targeting Facebook advertisements using personally relevant keywords also has the potential to decrease advertisements being shown to individuals who are not in the intended target population. Facebook also offers the opportunity to reach older individuals via their mobile device. In this study, most Facebook advertisements were seen in an individuals' mobile newsfeed (74.3%; n=42,059) indicating that many individuals aged 55 years and older are accessing Web-based content using their smartphones or tablet devices. Although Facebook advertising metrics did not allow for a breakdown by age and device type, this is a new trend that should be considered further for potential recruitment opportunities with this demographic and studied.

Facebook-targeted advertisement has great potential as a recruitment strategy for other studies. Previous studies using Facebook to recruit research participants have used a Facebook study group page to first recruit Facebook users to "like" and join the group page followed by direct advertisement to page group members via postings on the group page [29,30]. Our findings extend the work of Valdez et al [29] who assessed leveraging Facebook's social structure for research recruitment and concluded that Facebook recruitment was feasible for recruiting small samples for qualitative research but not for recruiting larger samples for quantitative research. Both recruitment strategies in the study by Valdez et al depended on an intermediary within the social structure of Facebook to recruit participants. For the first method tested, the study by Valdez et al created a study-related Facebook group page and asked administrators of other Facebook group pages to publicize the Valdez study-related Facebook group page to their members. The quantitative survey was then advertised on the study-related Facebook group page. The second recruitment method tested involved obtaining consent of other established Facebook group pages and posting a link to the study's survey on those pages. Our study did not use an intermediary component as the primary mode of recruitment, rather, we targeted keywords of individual profile "likes" and "interests" that the researchers felt were likely to appear in the profile of long-term current and former smokers (examples include "tobacco," "tobacco smoking," "smoking," "smoking cessation," "cigarettes," and "electronic cigarettes"). Most importantly, our findings extend the work of researchers who have demonstrated Facebook's utility with younger smokers [9,17,23,24] as we demonstrate Facebook is an effective recruitment tool for older smokers. Furthermore,

our findings extend the work of Frandsen et al [12] by evaluating the yield of the recruitment strategy by monitoring the time (in days) that the Facebook-targeted advertisement was in use, and clicks to website generated and sociodemographic characteristics of the participants recruited. In addition, leveraging creation of a Facebook group page that may appeal to the intended audience may be of benefit to recruitment efforts. To date, the Healthy Lungs Initiative Facebook group page has 114 likes in the 10 months since set up with very minimal effort. With a more concerted effort such as daily or weekly posting of general information to engage the Facebook group page audience, the audience network has the potential to grow. By attracting an audience that may have similar interests as a researcher's target population, this venue within Facebook offers the potential for longer-term participant engagement for future research opportunities.

Limitations and Unique Considerations

As with all studies, this study is not without limitations. Although Facebook recruited more participants at a lower cost, it is important to note that the study design is limited by inequitable comparison of recruitment methods. The Facebook advertisement was a nationwide recruitment campaign, whereas cost limited the study's newspaper advertisements to 1 high-readership newspaper in the Midwest United States. Although participant sociodemographic characteristics were largely similar, recruitment efficacy comparisons may be impacted. Future efforts to evaluate recruitment efficacy by Facebook versus newspaper advertisement should be designed to compare recruitment in the same geographic location. It is also important to note that although the newspaper advertisement had a 3-day reach of 230,742 individuals aged 55 years and older, it is highly probable that not all 230,742 individuals saw the advertisement given the nature of print media. Similarly, the 56,621 Facebook users who were likely exposed to the Facebook advertisement in their newsfeed have a probability that not all will have paid attention to and read the advertisement. However, Facebook advertisements requires an individual to click the advertisement when seen to take advantage of the opportunity being presented as opposed to print media, which can be read and returned to later to take advantage of the opportunity presented in the advertisement. As researchers design future recruitment strategies, it is important to consider how a recruitment advertisement is presented and consumed to maximize potential reach.

Participants recruited via Facebook were primarily Caucasian indicating that Facebook may not be the best recruitment method for older long-term current and former smokers from racially and ethnically diverse backgrounds. This potential limitation of sample diversity necessitates comparing the study outcomes by groups to ensure there are no significant differences. Therefore, the potential lack of representativeness of the study sample may be an important limitation of Facebook advertisement for recruitment. Future research involving lung cancer screening–eligible current and former smokers using Facebook-targeted advertisement as a recruitment method should monitor recruitment numbers of African-Americans and other racially and ethnically diverse groups closely. Although the sample recruited was predominately non-Hispanic Caucasian

men and women, Facebook's advertising program does offer the ability to target advertisement campaigns by zip code. Future studies could be designed to target ethnically diverse zip codes throughout a particular location to increase the numbers of participants from underrepresented populations. In addition, it is also important to note that the success of recruitment in this study may be directly related to its low burden (ie, 1-time survey directly accessed by a weblink with a $10 gift card incentive).

Finally, Facebook advertisement has the potential to recruit long-term smokers who may not be interested in responding to recruitment advertisements in high-traffic, highly visible areas or in methods that require the potential participant to contact the researcher directly. Facebook allows a degree of anonymity not present in these other forms of research recruitment, which may be successful with individuals who may have concerns related to stigma. However, it should be noted that Facebook could also enhance feelings of stigma because of a constant reminder of their smoking status being presented via a Facebook advertisement recruiting smokers. Future research is needed to explore potential differences in stigma levels in those screening-eligible smokers recruited via Facebook versus other methods.

Conclusions

Facebook-targeted advertisement offers the opportunity to reach a large number of potential study participants at a low cost while providing privacy and anonymity increasing the likelihood of recruiting a population that may experience stigma. Health behavior research related to lung cancer screening participation is important, and recruitment of current and former long-term smokers is critical to the success of this research. Smokers may experience perceived stigma, which may influence the success of recruitment by more traditional methods. Researchers who desire to reach individuals engaging in stigmatized behaviors (such as smoking) should consider Facebook as a viable option for study recruitment.

In addition, traditional recruitment strategies such as newspaper advertisement can be costly from a financial and personnel perspective. Recruitment of older current and former long-term smokers via Facebook-targeted advertisement offers a novel, efficient, efficacious, and cost-effective recruitment strategy that was successful in this cross-sectional, descriptive study.

Acknowledgments

Conflicts of Interest

None declared.

References

1. American Cancer Society. Cancer Facts & Figures 2015. Atlanta: American; 2015. URL: http://www.cancer.org/acs/groups/content/@editorial/documents/document/acspc-044552.pdf [accessed 2016-04-27] [WebCite Cache ID 6h5B7gyb6]

2. National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med 2011 Aug 4;365(5):395-409 [FREE Full text] [doi: 10.1056/NEJMoa1102873] [Medline: 21714641]

3. Wender R, Fontham Elizabeth T H, Barrera E, Colditz GA, Church TR, Ettinger DS, Wolf Andrew M D, et al. American Cancer Society lung cancer screening guidelines. CA Cancer J Clin 2013;63(2):107-117 [FREE Full text] [doi: 10.3322/caac.21172] [Medline: 23315954]

4. Moyer VA, U.S. Preventive Services Task Force. Screening for lung cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2014 Mar 4;160(5):330-338. [doi: 10.7326/M13-2771] [Medline: 24378917]

5. Carter-Harris L. Lung cancer stigma as a barrier to medical help-seeking behavior: Practice implications. J Am Assoc Nurse Pract 2015 May;27(5):240-245. [doi: 10.1002/2327-6924.12227] [Medline: 25736473]

6. Cataldo JK, Jahan TM, Pongquan VL. Lung cancer stigma, depression, and quality of life among ever and never smokers. Eur J Oncol Nurs 2012 Jul;16(3):264-269 [FREE Full text] [doi: 10.1016/j.ejon.2011.06.008] [Medline: 21803653]

7. Chapple A, Ziebland S, McPherson A. Stigma, shame, and blame experienced by patients with lung cancer: qualitative study. BMJ 2004 Jun 19;328(7454):1470 [FREE Full text] [doi: 10.1136/bmj.38111.639734.7C] [Medline: 15194599]

8. Shen MJ, Hamann HA, Thomas AJ, Ostroff JS. Association between patient-provider communication and lung cancer stigma. Support Care Cancer 2016 May;24(5):2093-2099. [doi: 10.1007/s00520-015-3014-0] [Medline: 26553030]

9. Ramo DE, Rodriguez Theresa M S, Chavez K, Sommer MJ, Prochaska JJ. Facebook Recruitment of Young Adult Smokers for a Cessation Trial: Methods, Metrics, and Lessons Learned. Internet Interv 2014 Apr;1(2):58-64. [doi: 10.1016/j.invent.2014.05.001] [Medline: 25045624]

10. Gold J, Pedrana AE, Sacks-Davis R, Hellard ME, Chang S, Howard S, et al. A systematic examination of the use of online social networking sites for sexual health promotion. BMC Public Health 2011;11:583 [FREE Full text] [doi: 10.1186/1471-2458-11-583] [Medline: 21777470]

11.   Nelson EJ, Hughes J, Oakes JM, Pankow JS, Kulasingam SL. Estimation of geographic variation in human papillomavirus vaccine uptake in men and women: an online survey using facebook recruitment. J Med Internet Res 2014;16(9):e198 [FREE Full text] [doi: 10.2196/jmir.3506] [Medline: 25231937]

12.   Frandsen M, Walters J, Ferguson SG. Exploring the viability of using online social media advertising as a recruitment method for smoking cessation clinical trials. Nicotine Tob Res 2014 Feb;16(2):247-251. [doi: 10.1093/ntr/ntt157] [Medline: 24127266]

13.   Batterham PJ. Recruitment of mental health survey participants using Internet advertising: content, characteristics and cost effectiveness. Int J Methods Psychiatr Res 2014 Jun;23(2):184-191. [doi: 10.1002/mpr.1421] [Medline: 24615785]

14.   Mychasiuk R, Benzies K. Facebook: an effective tool for participant retention in longitudinal research. Child Care Health Dev 2012 Sep;38(5):753-756. [doi: 10.1111/j.1365-2214.2011.01326.x] [Medline: 21985571]

15.   Jones L, Saksvig BI, Grieser M, Young DR. Recruiting adolescent girls into a follow-up study: benefits of using a social networking website. Contemp Clin Trials 2012 Mar;33(2):268-272 [FREE Full text] [doi: 10.1016/j.cct.2011.10.011] [Medline: 22101207]

16.   Altshuler AL, Gerns Storey Helen L, Prager SW. Exploring abortion attitudes of US adolescents and young adults using social media. Contraception 2015 Mar;91(3):226-233. [doi: 10.1016/j.contraception.2014.11.009] [Medline: 25537853]

17.   Villanti AC, Jacobs MA, Zawistowski G, Brookover J, Stanton CA, Graham AL. Impact of Baseline Assessment Modality on Enrollment and Retention in a Facebook Smoking Cessation Study. J Med Internet Res 2015;17(7):e179 [FREE Full text] [doi: 10.2196/jmir.4341] [Medline: 26183789]

18.   Pedersen ER, Helmuth ED, Marshall GN, Schell TL, PunKay M, Kurz J. Using facebook to recruit young adult veterans: online mental health research. JMIR Res Protoc 2015;4(2):e63 [FREE Full text] [doi: 10.2196/resprot.3996] [Medline: 26033209]

19.   Amon KL, Paxton K, Klineberg E, Riley L, Hawke C, Steinbeck K. Insights into Facebook Pages: an early adolescent health research study page targeted at parents. Int J Adolesc Med Health 2016 Feb 1;28(1):69-77. [doi: 10.1515/ijamh-2014-0074] [Medline: 25781667]

20.   Fazzino TL, Rose GL, Pollack SM, Helzer JE. Recruiting U.S. and Canadian college students via social media for participation in a web-based brief intervention study. J Stud Alcohol Drugs 2015 Jan;76(1):127-132 [FREE Full text] [Medline: 25486401]

21.   Nicholas A, Bailey JV, Stevenson F, Murray E. The Sexunzipped trial: young people's views of participating in an online randomized controlled trial. J Med Internet Res 2013;15(12):e276 [FREE Full text] [doi: 10.2196/jmir.2647] [Medline: 24334198]

22.   Hernandez-Romieu AC, Sullivan PS, Sanchez TH, Kelley CF, Peterson JL, Del RC, et al. The comparability of men who have sex with men recruited from venue-time-space sampling and facebook: a cohort study. JMIR Res Protoc 2014;3(3):e37 [FREE Full text] [doi: 10.2196/resprot.3342] [Medline: 25048694]

23.   Haines-Saah RJ, Kelly MT, Oliffe JL, Bottorff JL. Picture Me Smokefree: a qualitative study using social media and digital photography to engage young adults in tobacco reduction and cessation. J Med Internet Res 2015;17(1):e27 [FREE Full text] [doi: 10.2196/jmir.4061] [Medline: 25624064]

24.   Rait MA, Prochaska JJ, Rubinstein ML. Recruitment of adolescents for a smoking study: use of traditional strategies and social media. Transl Behav Med 2015 Sep;5(3):254-259 [FREE Full text] [doi: 10.1007/s13142-015-0312-5] [Medline: 26327930]

25.   PR Newswire. URL: http://www.prnewswire.com/news-releases/facebook-reports-first-quarter-2015-results-300070539. html [accessed 2016-01-05] [WebCite Cache ID 6eJl0DkRY]

26.   Duggan M, Ellison N, Lampe C, Lenhart A, Madden M. Pew Internet & American Life Project. 2014. Demographics of key social networking platforms - 2014 URL: http://www.pewinternet.org/files/2015/01/PI_SocialMediaUpdate20144.pdf [WebCite Cache ID 6eJObDN0v]

27.   Carter-Harris L, Slavin J, Monahan PO, Rawl SM. Development and psychometric evaluation of the lung cancer screening health belief scales. CANCER NURSING: An International Journal for Cancer Care 2016:NA-NA (forthcoming).

28.   Kramer J, Rubin A, Coster W, Helmuth E, Hermos J, Rosenbloom D, et al. Strategies to address participant misrepresentation for eligibility in Web-based research. Int J Methods Psychiatr Res 2014 Mar;23(1):120-129 [FREE Full text] [doi: 10.1002/mpr.1415] [Medline: 24431134]

29.   Valdez RS, Guterbock TM, Thompson MJ, Reilly JD, Menefee HK, Bennici MS, et al. Beyond traditional advertisements: leveraging Facebook's social structures for research recruitment. J Med Internet Res 2014;16(10):e243 [FREE Full text] [doi: 10.2196/jmir.3786] [Medline: 25348050]

30.   Heffner JL, Wyszynski CM, Comstock B, Mercer LD, Bricker J. Overcoming recruitment challenges of web-based interventions for tobacco use: the case of web-based acceptance and commitment therapy for smoking cessation. Addict Behav 2013 Oct;38(10):2473-2476 [FREE Full text] [doi: 10.1016/j.addbeh.2013.05.004] [Medline: 23770645]

Abbreviations

LDCT: low-dose computed tomography
REDCap: Research Electronic Data Capture

XSL•FO

RenderX

Original Paper

How Health Care Professionals Use Social Media to Create Virtual Communities: An Integrative Review

Kaye Rolls[1,2,3*], RN, BAppSc; Margaret Hansen[4*], RN, MSN, CNL, EdD; Debra Jackson[5,6,7*], RN, PhD; Doug Elliott[2*], RN, PhD

[1]Agency for Clinical Innovation, Intensive Care Coordination and Monitoring Unit, NSW Health Department, Chatswood, Australia

[2]Faculty of Health, University of Technology Sydney, Sydney, Australia

[3]Sydney Nursing School, University of Sydney, Sydney, Australia

[4]School of Nursing and Health Professions, University of San Francisco, San Francisco, CA, United States

[5]Oxford Brookes University, Faculty of Health & Life Sciences, Oxford, United Kingdom

[6]Oxford University Hospitals, Oxford Institute of Nursing & Allied Health Research, NHS Foundation Trust, Oxford, United Kingdom

[7]University of New England, School of Health, Armidale, Australia

[*]all authors contributed equally

Corresponding Author:
Kaye Rolls, RN, BAppSc
Agency for Clinical Innovation
Intensive Care Coordination and Monitoring Unit
NSW Health Department
Sage Building, Albert Avenue
Chatswood, 2067
Australia
Phone: 61 2 9464 4692
Fax: 61 2 9464 4728
Email: kaye.d.rolls@student.uts.edu.au

Abstract

Background: Prevailing health care structures and cultures restrict intraprofessional communication, inhibiting knowledge dissemination and impacting the translation of research into practice. Virtual communities may facilitate professional networking and knowledge sharing in and between health care disciplines.

Objectives: This study aimed to review the literature on the use of social media by health care professionals in developing virtual communities that facilitate professional networking, knowledge sharing, and evidence-informed practice.

Methods: An integrative literature review was conducted to identify research published between 1990 and 2015. Search strategies sourced electronic databases (PubMed, CINAHL), snowball references, and tables of contents of 3 journals. Papers that evaluated social media use by health care professionals (unless within an education framework) using any research design (except for research protocols or narrative reviews) were included. Standardized data extraction and quality assessment tools were used.

Results: Overall, 72 studies were included: 44 qualitative (including 2 ethnographies, 26 qualitative descriptive, and 1 Q-sort) and 20 mixed-methods studies, and 8 literature reviews. The most common methods of data collection were Web-based observation (n=39), surveys (n=23), interviews (n=11), focus groups (n=2), and diaries (n=1). Study quality was mixed. Social media studied included Listservs (n=22), Twitter (n=18), general social media (n=17), discussion forums (n=7), Web 2.0 (n=3), virtual community of practice (n=3), wiki (n=1), and Facebook (n=1). A range of health care professionals were sampled in the studies, including physicians (n=24), nurses (n=15), allied health professionals (n=14), followed by health care professionals in general (n=8), a multidisciplinary clinical specialty area (n=9), and midwives (n=2). Of 36 virtual communities, 31 were monodiscipline for a discrete clinical specialty. Population uptake by the target group ranged from 1.6% to 29% (n=4). Evaluation using related theories of "planned behavior" and the "technology acceptance model" (n=3) suggests that social media use is mediated by an individual's positive attitude toward and accessibility of the media, which is reinforced by credible peers. The most common reason to establish a virtual community was to create a forum where relevant specialty knowledge could be shared and professional issues discussed (n=17). Most members demonstrated low posting behaviors but more frequent reading or accessing behaviors. The most common Web-based activity was request for and supply of specialty-specific clinical information. This knowledge sharing is facilitated

XSL•FO
RenderX

by a Web-based culture of collectivism, reciprocity, and a respectful noncompetitive environment. Findings suggest that health care professionals view virtual communities as valuable knowledge portals for sourcing clinically relevant and quality information that enables them to make more informed practice decisions.

Conclusions: There is emerging evidence that health care professionals use social media to develop virtual communities to share domain knowledge. These virtual communities, however, currently reflect tribal behaviors of clinicians that may continue to limit knowledge sharing. Further research is required to evaluate the effects of social media on knowledge distribution in clinical practice and importantly whether patient outcomes are significantly improved.

Introduction

Although modern health care organizations are purported to be knowledge intensive [1], current management structures and work practices do not always facilitate development of intellectual and structural capital [2] or innovation uptake, leading to challenges for translating research into practice (TRIP) [3]. Contemporary organizational [1] and learning theories [4] highlight learning and behavior as being socially constructed and therefore influenced by social networks [5]. However, despite implementation of clinical network structures aimed at improving patient care and facilitating knowledge sharing between health care professionals and across organizational boundaries, bureaucratic, hierarchical, and intraprofessional barriers persist [6].

Information technology and the Internet have revolutionized communication to such an extent that humans can now communicate with colleagues anywhere at any time using social media platforms. Within the health care literature, there are however polarized views regarding the benefits and negative aspects of professional social media use [7,8]. Given this evolving technological environment and related continuing professional debate, the purpose of this paper was to review the literature on the use of social media by health care professionals for facilitating professional networking, knowledge sharing, and evidence-informed practice. Theoretical frameworks used to embed the use of social media in enabling collegial networking, knowledge sharing, and supporting evidence-informed practice are explored in the following section for context, before the focused literature search and review.

Background

Professional networking is a process of establishing a mutually beneficial relationship with other like-minded professionals [9]. For an organization, professionals networking between separate operational units promote knowledge flow and diffusion of innovations, potentially leading to improved professional performance [3]. Evolving views of learning including community of practice [4] and connectivism [10] highlight that professional development can be achieved through collective learning within social groups or networks. With the creation of Web-based communities, social media apps may facilitate this networking and professional development, enabling interactions between individuals regardless of time, space, or geography [11,12]. The interrelated concepts and frameworks are described

in the following section as background for exploring this topic area: diffusion of innovations, learning theories, evidence-based practice, knowledge management, and work in health care practice and social media.

Diffusion of Innovations

This theory describes how a novel idea, practice, or object is adopted by a particular social group or network [13]. In health, these innovations include new equipment, research findings, or practices. Rogers [13] demonstrated that although heterophilous communication (when individuals do not share common attributes such as values or socioeconomic status) increases access to novel ideas, for the vast majority of individuals, adoption of an innovation is dependent on homophilous communication (when individuals share common attributes). Five distinct types of individuals in a social group were identified. "Innovators" and "early adopters" are the first to adopt innovations with use mediated by a higher income and having greater access to novel information because of their broader, heterophilous social networks. The "early majority" are in turn influenced to adopt practices by observing use of and/or recommendation by early adopters. Finally, the "late majority" and "laggards" are the last to adopt because their communication channels are limited to those that share their views and experiences (homophilous) and are unlikely to be exposed to nonredundant knowledge or differing opinions [5,14].

Contemporary understanding of diffusion of innovations acknowledges that organizational or group factors also exert a powerful influence on individuals and the organization [15-17]. In particular, interconnectedness (connections between organizational members and units) and external orientation (organizational leaders with external networks) are both mediated by communication channels (networking internally or external to the organization) [13,16,17]. Individual innovators and early adopters with communication channels outside their everyday social and professional networks will learn more new information [18-20], although unless these individuals hold a central position within their local social network, it is unlikely this new knowledge will become embedded locally [14]. Credibility of intrapersonal channels (eg, peer to peer or opinion leader to professional) makes these channels more influential on adoption decisions [13,15,18,21,22]. Current social networks in health care organizations are generally homophilous with strong professional boundaries [23,24], which tend to control clinical practice [25].

Learning Theories

Current views of learning also highlight the importance of interaction or networking between individuals for learning and professional development. As a social learning theory, community of practice (CoP) positions learning as a fundamentally social behavior where individuals learn through their interactions and participation in the world [4]. Within a CoP, members acknowledge a shared knowledge domain (craft knowledge), practice, and identity [4]. Professional development therefore occurs during everyday workplace interactions, where important "how to" knowledge can only be gained from other colleagues [26]. For health care professionals, CoP is particularly relevant as the theory provides a framework for understanding the professional development of individuals within the workplace through different forms of participation [4,27,28]. At present, however, the effectiveness of health care CoPs to facilitate professional development and improve clinical practice needs further investigation because projects to date have operationalized and measured the effectiveness of the CoP in different ways [29,30].

Evidence-Based Practice

Recent literature on adoption of evidence-based practice [3] suggests that traditional health care structures do not create learning organizations that support: (1) development of intellectual capital [2]; (2) knowledge work [31]; or (3) assimilation of research findings into practice [32]. Furthermore, as knowledge does not flow freely between the silos of academia, clinical practice, publishing, and health care organizations, variations in the types and quality of care are common [33]. In health care, there have been mixed results where these channels (eg, opinion leaders) have been used to promote evidence-base practice [34,35] and peer-to-peer communication becomes more important as final adoption decisions are made [21]. In practice, however, clinicians continue to rely on personal knowledge (gained through education and experience) before seeking advice from close credible colleagues [36-39], despite the veracity of this advice not being critiqued or evaluated [36].

Knowledge Management, Knowledge Work, and Health Care Practice

Currently, organizational productivity [40], improved health outcomes, and cost-effectiveness are linked to the presence of a definitive knowledge management strategy that supports activities of "knowledge workers" [41]. Contemporary knowledge management strategies focus on human and contextual elements of knowledge, such as how knowledge is created and diffused through an organization [42,43]. Interorganizational and intraorganizational networks are central to knowledge creation and diffusion, given that much knowledge is experiential, implicit, or tacit [44], particularly in health care organizations.

Knowledge work involves evaluating data from novel situations and applying specialized and expertise transfer, to discover or create knowledge in a given context [45]. Health care professionals (nurses, physicians, and allied health disciplines) are a subgroup of knowledge workers identified as "technologists," where a personal knowledge store, initially based on formal academic education, evolves through experience and professional development [2]. Knowledge work can therefore be viewed as a form of evidence-based practice because it is the active thoughtful mode of work where clinicians decide how best to apply current knowledge, both personal and evidence, to individual patient care and other practice situations.

Social Media

Computers, the Internet, and social media have revolutionized human communication [46]. Web 1.0, existing between 1980 and 2000, was a Web-based environment characterized by static webpages with centralized creation, control, and distribution of content [47] and user interactivity facilitated by early social media (discussion forums, bulletin boards, and Listservs) [48]. The range of social media platforms exploded with arrival of Web 2.0, enabling new technologies including social and professional networking sites (eg, Facebook and LinkedIn), thematic networks, microblogs, wikis, social photo and video sharing tools, collaborative filtering tools, and multiuser virtual environments [49-51].

Aided by diffusion of tablet technology, Internet access, and improved mobile connectivity, use of social media has increased exponentially over the past few years. Between 2015 and 2016, both Internet and social media users increased by 10% to 46% (3.419 billion) and 31% (2.307 billion), respectively; there are however significant regional and national differences [52]. With respect to Internet use, Iceland has the highest penetration (98%) followed by Bermuda (97%) and Norway, Denmark, Andorra, and United Arab Emirates next (96%). North Korea has the lowest population usage (0.03%) followed by a number of Central African countries with less than 5%. Active population use of an social media account is greatest in North America (59%), South America (50%), East Asia, and Western Europe (48% each) and lowest in Central Asia (6%) and South Asia (11%). Social media use is similar across Western nations (eg, 58% Australia, 59% United States, 59% United Kingdom) but less in China at 47%. Although Facebook continues to dominate the social sphere, with 1590 million active accounts, users appear to be gravitating toward apps for networking including WhatsApp (900 million), QQ (860 million), and Facebook messenger (800 million). Among other platforms, Tumblr, Instagram, and Twitter continue to experience growth, whereas Skype and LinkedIn are stable [52]. For this paper, we adopted the International Medical Informatics Association's [51] classification, which identifies 13 types of social media platforms (see Textbox 1).

Textbox 1. Social media types [51].

- Social networks
- Professional networks
- Thematic networks
- Microblogs
- Blogs
- Wikis
- Forums or Listserv
- Social photo and video sharing tools
- Collaborative filtering tools
- Multiuser virtual environments
- Social apps and games
- Integration of social media with health information technologies
- Other (eg, FriendFeed)

Importantly, not all social media apps have the functionality to promote development of a Web-based or virtual professional community. The success of interactive conversational technologies (including discussion forums, Listservs, wikis, blogs, microblogs, and social networking sites (SNS), is contingent on members joining and participating in ongoing interaction; these are therefore the main types of social media platforms capable of creating virtual communities. Although virtual communities have been examined by a number of researchers from different disciplines, at this time, there is no universally accepted definition [53]. For this paper, we define a Web-based (virtual) community as "… a group of people who share a strong common interest, form relationships and interact online" [53] (p. 3). A community's existence depends on the structural capital produced from relationships established by member interaction and sharing of resources through the network [54]. Increasing numbers of organizations, professionals, and patients are now using social media to communicate and interact both internally and externally [55]. These real-life virtual communities or networks created by social media establish intrapersonal communication channels, overcoming barriers of time and geography, empowering users to communicate and interact (network) with a broad range of colleagues [11].

The purpose of this review was therefore to examine the research literature to identify how health care professionals use social media to develop virtual communities that facilitate professional networking, knowledge sharing, and evidence-informed practice. This review will add to the current literature by developing an understanding of how health care professionals use social media on a purely voluntary basis including integration of new media and behaviors such as conference tweeting.

Methods

Within the context of learning theories, diffusion of innovation, and social media in health care, an integrative literature review [56] was conducted to evaluate whether health care professionals

have been able to effectively leverage social media platforms to develop virtual professional communities that facilitate professional networking, knowledge sharing, and evidence-informed practice.

Literature Search

Two major electronic health databases, CINAHL and PubMed, were searched for research papers published between January 1990 and December 2015. Keywords were used as they applied to the main concepts of social media, networking, and professional development including virtual communities, social media, computer-mediated communication, Listserv, discussion forum, networking, Twitter, and Facebook. Additional search strategies included a review of reference lists of the papers and handsearching the table of contents of key journals (see Multimedia Appendix 1 for detailed search).

Papers that fulfilled the following criteria were selected for review: involved HCP participation exclusively as a voluntary activity; English language; peer-reviewed; and all research designs that highlighted HCP interaction using social media to develop a virtual community as the core component. Social media included were Listservs, discussion forums, SNS, and microblogs. Papers were excluded if they: (1) described a project within an education context including undergraduate or postgraduate learning or organizational education or training; (2) study protocol; and (3) narrative review. The first author extracted data from studies using a standard data extraction tool [57].

Study Methods Evaluation

After data extraction, the quality of each study was evaluated by 2 authors using standardized criteria. The CASP appraisal tool was used for qualitative studies (not including studies using content analysis) [58]. For studies using content analysis techniques [59-61], this included:

1. Data: appropriateness to research question, data corpus, sampling unit, unit of analysis, and sampling plan (described and justified).

2. Coding schema: appropriateness of approach, development, coders, training, theoretical underpinning of categories, and reliability of coding schema.

3. Analysis: appropriateness of approach.

Quality criteria for surveys included: (1) research question and design; (2) sampling framework and participant understanding; (3) instrument metrics; (4) response rate; (5) coding and analysis; and (6) result presentation [16]. The Scottish Intercollegiate Guidelines Network [62] appraisal tool was used for literature reviews. Studies were categorized as strong (most elements described with satisfactory quality), moderate, fair, and limited (poor reporting or description of research method).

Data Analysis

After data extraction and evaluation of study quality, summary tables were constructed to reduce data into manageable frameworks [56] and facilitate identification of patterns. These tables included data pertaining to:

1. Research overview including context, social media type, research design, sample and/or data corpus, data analysis and quality.

2. Web-based behavior including manifest and latent characteristics of emails or tweets and posting habits of members.

3. Reasons for belonging to a virtual community including meaning or value of community to members and motivators and barriers to Web-based participation.

4. Descriptions of Web-based communities including context, membership, and reasons or objectives for establishing Web-based community.

5. Research examining general social media use.

Only a limited amount of quantitative data could be aggregated for comparison across studies because of different data collection methods and outcomes. Qualitative data were synthesized to identify consistent patterns and themes.

Findings

Overall, 72 studies were included in the final review (see Figure 1 [63] and Multimedia Appendix 2). An overview of studies including context, design, instruments and data collection, sample and data corpus, data analysis, and study quality is listed in Multimedia Appendix 2). Findings are presented in the following sections:

1. Overview of research methods and critique of study quality

2. Social media use by health care professionals.

3. Web-based posting behaviors including the manifest and latent content of communication (emails, posts, or tweets) and posting habits.

4. Mediating factors of Web-based posting.

Figure 1. Literature search using PRISMA guidelines [63].
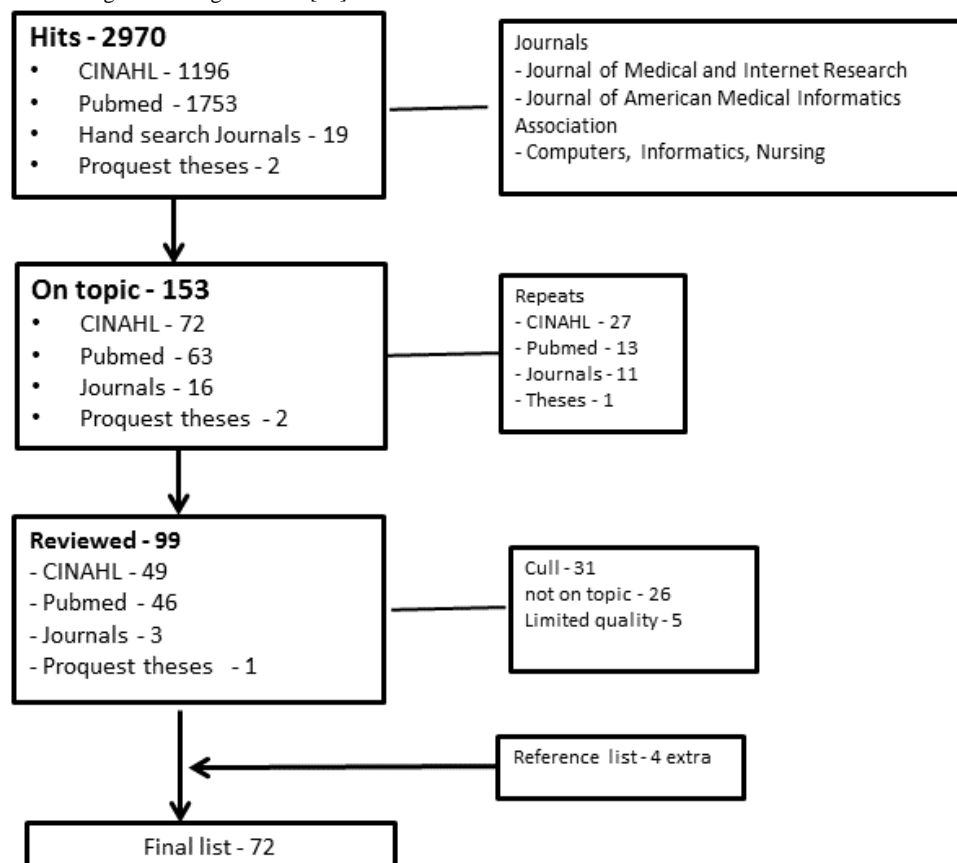
Overview of Research Methods and Critique of Study Quality

Of the 72 studies selected, there were 44 qualitative, 20 mixed methods, and 8 literature reviews. The most common methods of data collection were Web-based observation (n=30 studies), surveys (n=18), interviews (n=12), and focus groups (n=2).

Qualitative methods included: (1) qualitative (n=14; survey 11; discourse analysis 1; and interviews 2); (2) qualitative descriptive (n=26; content analysis 19, descriptive 5; thematic 1; social network analysis 1); (3) ethnography (n=2), Q-sort (n=1), and social network analysis (n=1). Q-sort is a multilevel study method where qualitative (subjective) responses are refined to develop a quantitative understanding or hierarchy of the phenomenon of interest [128]. Of the 20 mixed-method studies, combinations of methods included: content analysis and interviews (n=5); content analysis and survey (n=3); content analysis, survey, and social network analysis (n=1); Web-based observation and thematic analysis (n=1); Web-based observation and social network analysis (n=2); survey and diaries (n=1); survey and interviews (n=2); survey, interviews and observation (n=1); and survey and Web-based observation (n=2). Overall, the quality of these qualitative studies was satisfactory, with most fulfilling the CASP criteria [58] (see Multimedia Appendix 3). The quality of studies using content analysis (see Multimedia Appendix 4), survey methods (see Multimedia Appendix 5), or literature review (Multimedia Appendix 6) was mixed.

Content Analysis

Content analysis was commonly used in studies to reveal the content and meaning of textual data, which remains embedded in its origin or context [59]. In relation to Web-based communication, this approach can reveal the acquisition of new knowledge and skills and the social construction of knowledge [64]. A total of 30 studies used Web-based observation to collect emails, discussion threads, or tweets and applied either inductive (n=10) or deductive (n=20) content analysis techniques to identify: manifest content (topic, type of post, type of knowledge, frequency, discussion thread length, and/or participation rate) and latent content (accuracy of information, presence of knowledge work, or sophistication of discussion). Listservs or mailing lists (n=15) were the most common social media type examined followed Twitter (10), discussion forums (n=3), Web-based journal clubs (n=1), and Facebook (n=1).

The quality of studies was evaluated as high quality (n=12) or moderate quality (n=8), with the remaining 10 only fulfilling a limited number of required criteria (See Multimedia Appendix 4). Common study limitations affecting the validity of results included failure to report or justify the following elements: (1) data corpus and/or sampling unit [65-69]; (2) unit of analysis [65-67,69-74]; (3) coding schema development and categories with a limited theoretical basis for categories [66,68,69,73,75-77]; and (4) evaluate inter-rater reliability [66,68,69,71,73,78-80]. Only 2 studies kept the unit of analysis (that is post or email) within its contextual unit (ie, discussion thread) [77,81]. Sampling methods to gather the data corpus for analysis varied considerably. Most reports describe using a census sample [31,65,66,70,71,76,77,79,82-86] with stratified [27,80,87-91] or convenience samples [67-69,75,81,92,93] were

used less often. A random sample was used only once [74]; however, this was not adequately described.

Surveys

A survey design was used by 23 studies to examine member experiences with or intentions to use social media; only 2 demonstrated strong quality, 9 were moderate quality, and 12 were fair quality (see Multimedia Appendix 5). Methodological limitations impacting on the validity and generalizability of these findings included: (1) limited information regarding survey tool development [66,94-104] and (2) sampling bias including recruitment methods, low response rate, and/or failure to identify whether respondents were representative of community membership or study population [66,83,96,99,101-109].

Literature Reviews

Eight literature reviews were identified (4 systematic; 2 scoping; and 2 with no specific descriptor) with variable quality demonstrated (see Multimedia Appendix 6) [62]. The main deficits were: limited description of method [110,111]; a search strategy that was limited by years and/or databases [110-112]; and failure to evaluate the quality of studies covered [111-115]. Although each review had different questions, there were overlapping content areas: (1) social media adoption by clinicians [110], pharmacists [112,113], and radiographers [111]; (2) social media use for communication between patients, patient–clinician, or clinicians [116]; (3) type of social media use by clinicians [115]; (4) virtual communities for general practitioner professional development [12]; and (5) Twitter journal clubs [114]. Two studies [115,116] used the same definition of social media [117].

Overall, the quality of studies was mixed with 41 of moderate or higher methodological quality (strong 17; satisfactory 10; moderate 18) with 21 being of fair (17) or limited (4) quality, and there were 6 where we were unable to apply a quality framework. Despite a lack of methodological quality for a significant proportion of studies, all were retained in the review because of the limited contemporary evidence base and to therefore provide a comprehensive synthesis of this topic area.

Social Media Used by Health Care Professionals

Health care professionals currently use a broad range of social media platforms in practice, although understanding the extent is limited by study methods used and a lack of population data. Previous literature reviews [12,110-113,115,116] described use of most social media by most HCP groups to communicate interprofessionally and intraprofessionally and with health care consumers. The common types of social media platforms identified in this review were Listservs (n=22), Twitter (n=18), general social media (n=17), discussion forums (n=7), Web 2.0 (n=3), topic-specific discussion forums plus document repositories (n=3), a wiki (n=1), and Facebook (n=1). Physicians (n=24) in general and from 14 clinical specialties were the most common professional group studied, followed by nurses (n=15) in general and from 9 specialty areas, 4 groups of allied health professionals (n=14), health care professionals in general (n=8), a multidisciplinary clinical specialty area (n=9), and midwives (n=2).

Four papers described the uptake and use based on a population of potential users. Twenty percent or more had joined Listservs for occupational health practitioners [65], nurse practitioners [28], and intensive care (nurse data only) [118]. Although 209 of 1559 (13.07%) Korean emergency physicians had participated on a Facebook page [98], only 1.6% of US and 1.7% of Australian emergency physicians had joined Twitter by 2011 [119].

A number of studies of variable quality evaluated the general use of social media and found that health care professionals reported or demonstrated limited use of social media for professional purposes, and when they did, they preferred specialty-specific closed communities. Only 2 studies however were of a high-to-moderate quality [97,120]. A study examining US physicians' professional use of social media for connecting with colleagues reported limited use; only 52% currently used closed Web-based communities, 25% used wikis, whereas less than 20% used Facebook, podcasts, blogs, or Twitter. More than half also indicated they were unlikely to use these latter 4 platforms in the future [120]. A mixed-method study [97] used diaries to directly track the use of Web 2.0 by 35 junior physicians; 2.6 medical sites were accessed per day, and 52.9% of these visits were to user-generated platforms, but, there was limited professional use of SNS. A study of a broad range of Australian HCP found limited use and knowledge of Web 2.0 technologies; although the response rate was 9.2% (89/965), there were limited responses by physicians, and the researcher was unable to distribute to nurses [107]. The remaining surveys, of Greek health care professionals [109], pharmacy preceptors [106], mental health [101], family physicians [102], and urologists [99,104] found limited social media use, including social networks, for professional purposes. A single study of limited quality [103] found that 80.0% of respondents were using social media for professional purposes; however, the specific purpose was highly variable with only 44.1% using it for professional networking and 26.9% for obtaining or disseminating research evidence and professional development.

Two theories were applied across 3 studies to understand actual or future use of social media by health care professionals. Two high-quality studies applied the theory of planned behavior; a survey on the future use of Web 2.0 by Hong Kong nurses [121], and a qualitative study on the use of a wiki to transfer best practice care for patients with head injuries, where nurses were considered credible or influential peers by physicians [122]. Another survey of US physicians [120] applied the technology acceptance model (explains human behavior in relation to computer use) to explore user adoption. To use social media, clinicians required a positive attitude that the media was easy to use (usability), they were able to have a practice run to see how it worked (trialability), the platform worked better than current solutions (relative advantage), and the technology was accessible in the workplace and fitted in with current work practices. The final mediating factor was that their peers also shared these attitudes, a reflection of the influence of homophily.

Social Media and Virtual Communities

Overall, 36 reports described 31 discrete virtual communities [27,31,65-69,71,74,76,79,81-84,88,92,94-96,98,100,105,118,123-126]

that were established in 3 main ways. The most common were discussion forums or Listservs created by a professional society [28,65-67,69-71,74,76,79,83,91,94,98,105,123]. Nine communities appear to be have been established by an individual or group of HCP using inexpensive or open access platforms such as Yahoo groups, mailing list software, or Twitter [27,68,81,84,88,92,95,124,127]. Eight communities were established by a government health department with the purpose of improving communication and knowledge distribution between health care professionals to enhance care [31,82,96,100,118, 125,126,128,129].

The most common reasons for establishing a discrete virtual community were to:

1. Create a professional forum where relevant professional and academic issues could be discussed and information and knowledge shared [28,67,69,70,73,76,79,91,95,98,100, 105,125,127-130].

2. Address professional isolation [70,73,91,105, 126-128].

3. Facilitate networking [27,76,83,91,100,105,124,127,128].

4. Foster peer collaboration and mentoring [83,100,105,128].

5. Facilitate professional development [74,83,91].

6. Improve clinical practice through research and evidence translation [100,127].

7. Obtain clinical advice or opinion [98].

Where a distinct professional community was evaluated, 31 of 36 were a virtual community in a single HCP discipline such as physician, nurse, occupational therapist, social worker, pharmacist, or medical librarian. Note that these virtual communities were for a specific clinical specialty, except for 2 nursing communities (Nursenet [81,92] and allnurses [68]) and the medical librarian virtual community [73,94]. Five multidisciplinary virtual communities were all Listservs for a clinical specialty established by: (1) an international professional society for travel medicine clinicians [71]; (2) a Norwegian professional society for occupational health clinicians [65]; (3) Spanish speaking radiation medicine clinicians [66]; (4) an Australian jurisdiction–based health unit for intensive care clinicians [118,126] and a Twitter network connecting physicians from 3 specialty areas [131].

Social network analysis of 3 virtual CoPs demonstrated early evidence supporting the flow of knowledge across virtual communities. A study examining the growth and social network of an intensive care Listserv demonstrated an evolution from a single-state nurse-specific network to an Australian-wide, multidisciplinary, and multiorganizational network over 6 years [118]. A distinct Twitter virtual community, created via following patterns of emergency physicians (board certified in United States or Australia) showed a small core (2.8%) with a larger interconnected group, although 34% were not connected to any others [119]. Another study examined Twitter virtual community connections across 4 physician groups from the United States and reported 4 distinct communities with a small overlap where there was some information flow between groups [131].

XSL·FO

RenderX

The question of whether a CoP might be possible using social media was examined in several studies. Three high-quality qualitative studies exploring a critical care nursing Listserv found that motivators of Web-based knowledge sharing mapped to key aspects of CoP theory, including reciprocity, collectivism, respectful environment, and altruism [27,88,124]. A survey of a nurse practitioner Listserv reported that a sense of community correlated with learning (Pearson coefficient = +.94) and connectedness (r=.95), although the response rate was only 22%, and there was no indication whether respondents were active posters or nonposters [28]. A literature review [12] adapted a 7-item framework for a health care virtual CoP from a business model [132], exploring: (1) facilitation; (2) champion and support; (3) objectives and goals; (4) a broad church; (5)

supportive environment; (6) measurement, benchmarking, and feedback; and (7) technology and community.

How Members Use Social Media Virtual Communities

Most research on how health care virtual communities were used by members focused on posting behaviors. Web-based roles of members can be broadly described as participants (Web-based posters) and nonposters. Direct measurement of posting behaviors across a number of platforms demonstrated a pattern of a minority of members being responsible for most posts [31,65-67,69,82,94,130] or conference tweeting [77,86,87,89,90] (see Table 1; Multimedia Appendix 6). The same pattern was revealed across 4 surveys asking health care professionals about their Web-based behavior [69,94,96,126].

Table 1. Summary of studies examining Web-based posting behaviors by virtual community members.

Reference	Social media; time span	Nonposting	Low posting	Medium posting	High posting
Cervantez Thompson [76]	Listserv; 18 months	33%	27.8% at least once in 18 months		10 members > 30
Long et al [70]	Discussion forum; 12 months	28.3% (n=170)	48% (n=239) < 4 times	30% (n=179) 4-20 times	0.2% (n=12) 19-59 times (17% of total data corpus)
Stewart et al [125]	Discussion forum; 27 months	33% (n=14)	46% (n=21) < 14 times	13% (b=6) 15-28 times	9% (n=5) 29-56 times
Rodriguez-Recio and Sendra-Portero [66]	Listserv; 5 years	46.3% (n=175)	434% (n=161) 1-10 times	8% (n=30) 11-30 times	3.2% (n=12) 31 to < 200 times
Macdonald et al [71]	Listserv; 6 months				Top 20 users—43% of posts
Morken et al [65]	Listserv; 1997-2004		Average number of posts = 2.1; this reduced to 0.6 in 2004		
Brooks and Scott [31,82]	Discussion forum		11 aged care nurses posted over 7 months	26 cardiac nurses posted over 7 months	29 midwives posted over 1.5 months

Conversely, "non-posting" or "lurking" behavior [133] was generally high, ranging from 28% to 46% (see Table 2). These findings however do not indicate whether nonposters were active in reading posts. Where being active nonposters was directly measured, it ranged from 1% to 33%, whereas survey respondents self-reported reading levels of post as 64% to 96% (see Table 2).

Table 2. Summary of studies examining reading (access) behaviors.

Reading	Social media	0	Low	Medium	High
Stewart and Abidi [125]	Discussion forum; Web-based observation; access	32.6%	54.3%	4.3%	8.7%
Cook-Craig and Sabah [96]	Discussion forum; Web-based observation; access	1%	11%	38%	50%
Rolls et al [126]	Listserv; survey		3.5%	13.2%	83%
Schoch and Shooshan [94]	Listserv; survey; access		36%	24%	40%
Kim et al [98]	Facebook; survey; access	once or less each week—22.3	2-4 times per week—23.7%	5-6 times per week—16.6%	> 1 per day—37.4%
Whitaker et al [69]	Listserv; survey	Seldom or never 10%	1 per week to month	Several times per week 40%	Daily 40%

Current evidence describing barriers and motivators to posting over the Internet is difficult to quantify; only 4 studies examined these elements, 2 of which reviewed the same Listserv and

included frequent poster activity [27,88] (see Table 3). These limited data suggest a symbiotic relationship between members and the Web-based community, with behaviors of posters

influenced by both access to new knowledge and contributing for other members of the community. These elements of altruism, reciprocity, and collectivism are essential components of CoP building [27,28,88]. Reported barriers suggest that knowledge self-efficacy and time are key mediators of Web-based participation or knowledge sharing in health care virtual communities [88,126].

Table 3. Mediators of posting in by health care professionals in social media and virtual communities.

	Individual level	Community level
Motivators	Personal gain: (1) more knowledge [28,77,88]; (2) a better reputation [88]; (3) emotional support [88]	Collectivism [28,77,88,124]
	Seeker interest [88]	Reciprocity [28,88,124]
	Altruism [28,77,88,124]	Respectful environment [88,124]
	Self-selection [27]	Technology [88,124]
	Validation of one's practice [27]	Asynchronous nature [27]
	Advocacy [77]	Facilitate networking [77]
	Better understanding of current knowledge and best practice in the field [27,77,103]	Noncompetitive environment [27]
Barriers	Nothing to add [88]	Information not trustworthy [103]
	Nothing to say [126]	Lack of privacy [28,103]
	Lack of time [88,103,126]	Technology [88,103]
	Unfamiliarity with subject [88]	Confidentiality of sharing organization documents [88]
	Lack of confidence [126]	Tone of discussion [28,126]
	Local unit constraints [126]	Alienation[28]
	Attitude of seeker [88] or poster agenda[103]	Unprofessional behavior [28,103]

Overall, these findings supported the use of social media by health care professionals, specifically discussion forums and mailing lists platforms, to develop virtual professional CoPs. These communities valued the Web-based forums as information or knowledge portals, enabling members to "keep up to date" [73,94,124] with clinically relevant and quality information [66], develop workplace resources [123] and benchmark practice [27,123,124]. Importantly, access to a broader range of professional colleagues beyond their local organization enabled members to make more informed practice decisions, with greater confidence that these decisions reflected current best practice [124].

Manifest Content of Posts

Manifest content is the text immediately visible and easy to identify and count [61]. The quality of evidence describing the manifest content of posts, including posting behaviors, number of posts, length of discussion thread, and ratio of initial post to responses, was limited by both the quality of studies (See Multimedia Appendix 7) and variability in the sampling and measurement methods used. Making sense of the types of posts in social media was also challenging as researchers used variable descriptors when categorizing post types. The proportion of clinical versus nonclinical posts varied greatly across studies. Clinical posts were in the majority across 5 Listservs: travel medicine professionals (88%) [71]; radiology professionals (71.8%) [66]; rehabilitation nurses (60%) [76]; forensic occupational therapists (59.9%) [127]; and occupational health (54%) [65]. Posts on professional issues were more common on a plastic surgery discussion forum (60% concerned education and introduction of new members) [79] and an international nursing discussion forum (83% focused on career and education advice, work issues, and handling job-related emotions) [68]. Analyzing categories of conference tweets revealed similar results to Listserv and discussion forum data; however, understanding how it related to clinical knowledge or new research was difficult because of variable taxonomies and mixed quality. Five studies, evaluating 8 conferences, used the same taxonomy [134] and found that tweets concerning conference content (termed informative) ranged between 20% and 30% [78,87], 30% and 40% [78,135], 40% and 50% [78,136], and 50% and 60% [78]. Similar data were found across 2 conference years where most tweets from an oncology conference were clinical topics (54.5% and 60.4%), such as clinical management discussions and clinical news or trial outcome [80]. Contrasted against this was a study of an emergency conference, which found that 75% of tweets related to conference content [90]. Note however that the most commonly used taxonomy [134] has limited validity within or generalizability to health care conference data, as it was developed from a single Twitter feed specific to the author, it was not reviewed by a second coder or tested against another dataset. A systematic review of Twitter journal clubs that cross-referenced hashtag use with Web-based data [114] found sustained and increasing use of 5 specific tags (#ADC_JC; #ebnjc; #IGSJC; #Nephjc; and #urojc).

Four studies of mixed quality found that topics of clinical posts in virtual communities mapped to the knowledge domain of a professional specialty. Within a travel medicine Listserv, there were 27 topics across 5 major categories (vaccine preventable diseases, vector-borne diseases, pretravel, general, and miscellaneous) [71]. Pediatric occupational therapists posted on 4 categories (practice, performance component, performance area, and health conditions) [70]. Members of an occupational health forum posted on 4 clinical categories (chemical hazards, methods in health and safety environment, ergonomics, and noise and radiation) [65]. Pharmacists discussed a broad range of topics including patient and clinical problems, pharmacy politics, legal issues, drug tariffs, government policy, business and finance, risk management, and pharmacy information technology [69].

Latent Content of Posts

Latent content reflects the hidden meaning of textual content by a researcher [137]. Latent content examined included types of knowledge exchanged and presence of discussion and existence of knowledge work. Understanding the types of posts was limited by variability in study methods and challenging because of widely varying definitions and lack of robustly developed content analysis tools. Only 3 studies examined the types of knowledge within virtual community posts (Multimedia Appendix 6). Two high-quality studies that examined a nursing Listserv found that more than 90% of knowledge exchanged was practical knowledge (related to institutional practices, personal opinion, or suggestion) rather than book knowledge (facts, general regulations, statutes, or published works) [27,88]. On a Spanish radiological Listserv, 43% of emails were classified as scientific information [66].

As described earlier, knowledge work involves elements of interaction, critical reflection, and learning as a dialogical process [31,138]. Only limited data were identified supporting the presence of knowledge work within virtual professional communities. Three studies [66,73,130] described the presence of discussion or meta-discussion within emails exchanged; however, no content analysis tool or definitions were provided to justify these conclusions. One single high-quality study [81] effectively described the presence of reflection in discussion, where participants reported changes in practice through an iterative process that included off-line and Web-based discussions. One organizational project demonstrated mixed results, with high levels of knowledge work on a midwifery forum but lower levels in both aged care and cardiology forums [31,82,139].

Discussion

The focus of this review was to identify whether health care professionals have effectively created virtual communities to facilitate professional networking, knowledge sharing, and evidence-informed practice. The current evidence is mixed in terms of quality and type of studies undertaken. Apart from a couple of exceptions, studies published before 2004 were limited by common methodological limitations including sample and measurement bias, especially when content analysis techniques or surveys were used. The quality of more recent studies, including those using focus groups, surveys, interviews and Q-sort, has improved and reveals important insights into how health care professionals use social media to develop virtual communities and interact with professional colleagues. Importantly, these insights indicate that virtual communities may provide significant opportunities to overcome current barriers to knowledge flow and professional networking in health care.

This beginning evidence supports the view that health care professionals have adopted social media to create viable virtual professional communities, and that health care virtual communities share similar characteristics to other professional communities. A consistent pattern in Web-based communities was that most contributions were attributed to a limited number of individuals [31,65-67,71,76,82,130]. The voluntary nature of participation within social networks and virtual communities means that members participate at different levels and may adopt specific Web-based roles [140]. A virtual community is likely to have a mixture of lurkers, observers, passive, and active contributors [141]. Importantly, nonposting virtual community members continue to belong because of potential access to important information (reflective of Burnett's information neighborhood) [142], but, this requires further investigation.

There is a modest level of evidence that the most common activity in health care virtual communities is the exchange of experiential domain-specific knowledge. Importantly, the rise of conference tweeting and journal clubs suggests that Twitter may have a role in reducing the evidence practice gap. There are however only limited contemporary data supporting the transfer of empirical knowledge or how this new knowledge is used in practice [27,80,87,123,124]. In addition, although there are generally positive attitudes toward and intention to use social media [120,121], a skepticism persists regarding the veracity of information [97,103,122]. Understanding the exchange of knowledge remains limited as all but one study [77] failed to appreciate that social media interactions reflect a conversation with each post likely influenced by an antecedent [143].

Gaining access to previously unknown information or knowledge is an essential benefit of networking [20], and sharing this information is a major driver of social networks and virtual communities [144]. Effective knowledge transfer and innovation development occurs in social networks where there is a shared understanding of knowledge but also a density of ties providing access to novel information [20]. The symbiotic relationship between the culture of a virtual community and its members creates an ethos of knowledge sharing in a Web-based context. Similar to nonhealth virtual community [145,146], Web-based knowledge sharing is facilitated by a culture of altruism, trust, collectivism and reciprocity, as well as a respectful noncompetitive environment [28,88,124,126]. Knowledge self-efficacy, a belief the answer supplied is correct and worthwhile, influences knowledge sharing by individuals [147-150]. Moreover, group behaviors perceived as negative (eg, tone of discussion or contentious issues) have an undesirable effect on both willingness to share knowledge and retention of community members [28,126,142].

The dominance of Listservs and discussion forums in this search period is not surprising, given these platforms have been available since the early 1990s [48]. Although these social media platforms provide HCP with the ability to interact, they are limited in functionality, particularly with their capacity to create and/or store permanent community artifacts (such as guidelines or learning packages) required by a CoP for knowledge and practice development [4]. The relatively recent arrival of Web 2.0 platforms, enabling users to create and/or upload content, overcomes these problems [47]; however, there were only 2 reports [105,128] of virtual communities using this modality evident in this review. Conference tweeting, tweet chats, and journal clubs haves emerged in recent years; however, the current variability in methods used limits our understanding how this might contribute to distribution of scientific knowledge.

At this time, the evidence suggests that clinicians prefer to use social media that allows them to communicate within their own profession and within a clinical specialty, as most virtual communities identified were for a clinical specialty within a single HCP discipline. Although this may reflect continuing tribal behavior of clinicians in practice [23,24,151], monodiscipline social networks can create strong boundaries that inhibit interprofessional learning and knowledge sharing [152] and promoting practice initiatives to improve patient outcomes [151]. Sharing knowledge and adoption of innovation is enhanced where there is homophily (shared within a multidisciplinary clinical specialty domain such as emergency or intensive care) and credibility [152]. Because patients are commonly cared for by a multidisciplinary team and these clinicians generally share a common specialty knowledge domain, multidisciplinary networks are more likely to be effective in knowledge transfer and creation [20,42]. In this review, this potential was demonstrated in 2 multidisciplinary virtual communities [118,125]. A social medium that creates an open virtual community through user-generated follow patterns (such as Twitter) has this potential, but this is yet to be demonstrated in health care.

Strengths and Limitations of the Review

The key strengths of this review were the timeline, promoting the inclusion of the broad range of current social media apps, and the specific focus on voluntary professional participation. Previous reviews were unable to provide clear information on our focused question because of inclusion of education and undergraduates [110,115] or patients [116]. Nonetheless, exclusion of research within a training framework remains a limitation as does the exclusion of wikis and other collaborative writing technologies and blogs. Another limitation was the keyword search, where we were dependent on how keywords were applied when papers were published. Of note, the term social media was only added to the MeSH list in PubMed in 2012. We attempted to address this by undertaking a series of searches (see Multimedia Appendix 1) using a range of keywords; however, we may not have captured all relevant publications. Moreover, we only used English language publications, so we may have missed other important studies.

Recommendations for Further Research

As the current evidence is limited in quality and with most studies examining older technological platforms, there are a number of recommendations for future research. Recent studies [65,70,71] show that solicitation and supply of knowledge of craft-specific knowledge are the most common posts exchanged on professional health care virtual communities. There are limited data however to describe: (1) the specific types of knowledge exchanged (eg, scientific vs experiential or tacit vs explicit); (2) accuracy of this knowledge; (3) whether the knowledge supplied addressed what the poster requested; and (4) what the receivers of the emails, including the original poster, did with this knowledge. Further content analysis of posts using a more systematic approach may reveal not only the knowledge needs of members but also the knowledge embodied within the network.

At present, there is limited understanding of why individuals join or participate in a Web-based community; previous studies have generally examined activity from the perspective of Web-based posters. Some data suggest that professionals will join a virtual community where they find local resources inadequate [153]. Importantly, although nonposters or limited posters constitute a large portion of virtual community membership, it is not clear why they belong to the community or why they chose to limit posting. Because movement of knowledge or innovation into and around an organization is the role of boundary spanners and knowledge brokers (eg, educators or researchers), do these individuals see membership as a valuable tool for their substantive position, as preliminary data suggest [123]? If so, could health care organizations improve knowledge flow by facilitating communication between key personnel using Web-based communities? Understanding these phenomena is important if leaders or moderators of virtual communities, researchers, or health system change agents are to create optimal Web-based experiences and ensure the viability of the social medium within professional health care environments.

Early research suggested that Web-based forums may facilitate the development of higher order cognitive skills, such as tertiary students' critical thinking [154]. These important findings may be linked to educational design, implementation, and evaluation for effective adult learning by today's HCP. This contrasts with the self-selective and voluntary nature of professional forum membership. Only 2 studies verified the presence of a CoP within a Web-based health care community [28,124]. There is however now a worldwide education movement based around the use of social media for the professional development of clinicians. Free Web-based medical education (#FOAMed) [155] is an egalitarian movement promoting open Web-based publication of a wide range of resources for the education of any clinician. Further research is required however to identify the viability of social media platforms for voluntary professional development of health care professionals. This may require a mixed-methods approach to comprehensively understand the learning interaction (via a social network analysis), process (via content analysis), and outcome (via a survey) [156].

XSL•FO

RenderX

Conclusion

The current evidence on the use of social media by health care professionals suggests that virtual communities are viewed as valuable knowledge portals where craft knowledge is exchanged. This review, apart from the recent emergence of conference tweeting and Twitter journal clubs, found only a limited number of publications concerning newer social media platforms. Arguably, the current range of social media platforms and electronic devices facilitating exchange of information makes professional networking possible wherever the Internet is available. Given that a number of the current challenges of TRIP are related to a lack of inter professional and intraprofessional communication channels, there is significant potential within multidisciplinary virtual communities to facilitate the transfer of experiential and research knowledge by breaking down professional and organizational boundaries. Further research is required to evaluate whether virtual communities may improve patient outcomes by facilitating professional development, evidence-based practice, and elimination of clinical practice silos.

Authors' Contributions

KR, DE, MH, and DJ contributed to study design and manuscript preparation. KR and DJ contributed to data collection and analysis.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Detailed search strategy and results.

[PDF File (Adobe PDF File), 121KB - jmir_v18i6e166_app1.pdf]

Multimedia Appendix 2

Overview of all studies included in final review.

[PDF File (Adobe PDF File), 488KB - jmir_v18i6e166_app2.pdf]

Multimedia Appendix 3

Quality assessment table for qualitative studies.

[PDF File (Adobe PDF File), 288KB - jmir_v18i6e166_app3.pdf]

Multimedia Appendix 4

Quality assessment of studies using content analysis techniques.

[PDF File (Adobe PDF File), 400KB - jmir_v18i6e166_app4.pdf]

Multimedia Appendix 5

Quality checklist for questionnaire surveys.

[PDF File (Adobe PDF File), 439KB - jmir_v18i6e166_app5.pdf]

Multimedia Appendix 6

Quality assessment of reviews.

[PDF File (Adobe PDF File), 263KB - jmir_v18i6e166_app6.pdf]

Multimedia Appendix 7

Content of posts on health care social media.

[PDF File (Adobe PDF File), 446KB - jmir_v18i6e166_app7.pdf]

References

1. Drucker PF. The new society of organizations. Harv Bus Rev 1992;70(5):95-104. [Medline: 10121319]
2. Moody RC. Nurse productivity measures for the 21st century. Health Care Manage Rev 2004;29(2):98-106. [Medline: 15192982]
3. Oborn E, Barrett MG, Racko G. Working Paper Series. Cambridge: Cambridge Judge Business School, University of Cambridge; 2010. Knowledge translation in healthcare: a review of the literature URL: https://www.jbs.cam.ac.uk/fileadmin/user_upload/research/workingpapers/wp1005.pdf [WebCite Cache ID 6cU9Kj12O]
4. Wenger E. Communities of practicesocial learning systems. In: Starkey K, Tempest S, McKinlay A, editors. How organizations learn: managing the search for knowledge. London: Thomson; 2004:238-258.
5. Borgatti SP, Halgin DS. On Network Theory. Organization Science 2011 Oct;22(5):1168-1181. [doi: 10.1287/orsc.1100.0641]
6. Braithwaite J. Between-group behaviour in health care: gaps, edges, boundaries, disconnections, weak ties, spaces and holes. A systematic review. BMC Health Serv Res 2010;10:330 [FREE Full text] [doi: 10.1186/1472-6963-10-330] [Medline: 21134295]
7. Ferguson C. It's time for the nursing profession to leverage social media. J Adv Nurs 2013 Apr;69(4):745-747. [doi: 10.1111/jan.12036] [Medline: 23488814]
8. DeCamp M, Cunningham AM. Social media: the way forward or a waste of time for physicians? J R Coll Physicians Edinb 2013;43(4):318-322. [doi: 10.4997/JRCPE.2013.411] [Medline: 24350316]
9. Moore B. The Australian concise Oxford dictionary, 4th edition. South Melbourne, Victoria: Oxford University Press; 2004.
10. Siemens G. ITFORUM. 2008. Learning and Knowing in Networks: Changing roles for Educators and Designers URL: http://itforum.coe.uga.edu/Paper105/Siemens.pdf [accessed 2015-10-25] [WebCite Cache ID 6cWn7yESy]
11. Deng L, Poole M. Learning through ICT-enabled social networks. IJITM 2008;7(4):374-391. [doi: 10.1504/IJITM.2008.018655]
12. Barnett S, Jones SC, Bennett S, Iverson D, Bonney A. General practice training and virtual communities of practice - a review of the literature. BMC Fam Pract 2012;13:87 [FREE Full text] [doi: 10.1186/1471-2296-13-87] [Medline: 22905827]
13. Rogers E. Diffusion of innovations, 5th edition. New York: Free Press; 2003.
14. Tasselli S. Social networks of professionals in health care organizations: a review. Med Care Res Rev 2014 Dec;71(6):619-660. [doi: 10.1177/1077558714557079] [Medline: 25380607]
15. Fleuren M, Wiefferink K, Paulussen T. Determinants of innovation within health care organizations: literature review and Delphi study. Int J Qual Health Care 2004 Apr;16(2):107-123 [FREE Full text] [doi: 10.1093/intqhc/mzh030] [Medline: 15051705]
16. Kyriakidou O, Bate P, Peacock R, Greenhalgh T, MacFarlane F. Diffusion of innovations in health service organisations: a systematic literature review. Malden, Mass: Blackwell; 2005.
17. Grossan M, Apaydin M. A multi-dimensional framework of organizational innovation: a systematic review of the literature. Journal of Management Studies 2010;47(6):1154-1191. [doi: 10.1111/j.1467-6486.2009.00880]
18. Granovetter M. The strength of weak ties: a network theory revisited. In: Marsden P, Lin N, editors. Social Structure and Network Analysis. New York: John Wiley and Sons; 1982:105-130.
19. Singh T, Cullinane J. Social networks and marketing: potential and pitfalls. IJEMR 2010;3(3):202-220. [doi: 10.1504/IJEMR.2010.034829]
20. Nieves J, Osorio J. The role of social networks in knowledge creation. Knowledge Management Research & Practice 2013;11(1):62-77. [doi: 10.1057/kmrp.2012.28]
21. Locock L, Dopson S, Chambers D, Gabbay J. Understanding the role of opinion leaders in improving clinical effectiveness. Soc Sci Med 2001 Sep;53(6):745-757. [Medline: 11511050]
22. Valente T. Diffusion of innovations and policy decision-making. Journal of Communication 1993;43(1):30-45.
23. Creswick N, Westbrook JI, Braithwaite J. Understanding communication networks in the emergency department. BMC Health Serv Res 2009;9:247 [FREE Full text] [doi: 10.1186/1472-6963-9-247] [Medline: 20043845]
24. Creswick N, Westbrook JI. The medication advice-seeking network of staff in an Australian hospital renal ward. Stud Health Technol Inform 2007;130:217-231.
25. Duncan EM, Cuthbertson BH, Prior ME, Marshall AP, Wells EC, Todd LE, SuDDICU International Study Group. The views of health care professionals about selective decontamination of the digestive tract: an international, theoretically informed interview study. J Crit Care 2014 Aug;29(4):634-640. [doi: 10.1016/j.jcrc.2014.03.013] [Medline: 24747038]
26. Brown J, Duguid P. Organizational learning and communities of practice: Toward a unified view of working, learning and innovation. Organization Science 1991;2(1):40-57.
27. Hara N, Foon Hew K. Knowledge‐sharing in an online community of health‐care professionals. Info Technology & People 2007 Aug 28;20(3):235-261. [doi: 10.1108/09593840710822859]
28. Widemark E. Community and Learning: A Virtual Community of Practice for Nurse Practitioners, in Education. Ann Arbour, US: Capellea University; 2008.

29.   Ranmuthugala G, Plumb JJ, Cunningham FC, Georgiou A, Westbrook JI, Braithwaite J. How and why are communities of practice established in the healthcare sector? A systematic review of the literature. BMC Health Serv Res 2011;11:273 [FREE Full text] [doi: 10.1186/1472-6963-11-273] [Medline: 21999305]

30.   Li L, Grimshaw J, Neilsen C, Judd M, Coyte P, Graham I. Use of communities of practice in business and health care sectors: a systematic review. Implementation Science. 2009. p. 27 URL: http://www.implementationscience.com/content/4/1/27 [WebCite Cache ID 6iEsp86gk]

31.   Brooks F, Scott P. Knowledge work in nursing and midwifery: an evaluation through computer-mediated communication. Int J Nurs Stud 2006 Jan;43(1):83-97. [doi: 10.1016/j.ijnurstu.2005.02.003] [Medline: 16326164]

32.   Crow G. Diffusion of innovation: the leaders' role in creating the organizational context for evidence-based practice. Nurs Adm Q 2006;30(3):236-242. [Medline: 16878009]

33.   McGowan B. #SocialQI: simple solutions for improving your healthcare, 1st edition. USA: No Limit Publishing; 2012.

34.   Thompson GN, Estabrooks CA, Degner LF. Clarifying the concepts in knowledge transfer: a literature review. J Adv Nurs 2006 Mar;53(6):691-701. [doi: 10.1111/j.1365-2648.2006.03775.x] [Medline: 16553677]

35.   Doumit G, Gattellari M, Grimshaw J, O'Brien MA. Local opinion leaders: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2007(1):CD000125. [doi: 10.1002/14651858.CD000125.pub3] [Medline: 17253445]

36.   Marshall AP, West SH, Aitken LM. Preferred information sources for clinical decision making: critical care nurses' perceptions of information accessibility and usefulness. Worldviews Evid Based Nurs 2011 Dec;8(4):224-235. [doi: 10.1111/j.1741-6787.2011.00221.x] [Medline: 21649854]

37.   Spenceley SM, O'Leary KA, Chizawsky Lesa L K, Ross AJ, Estabrooks CA. Sources of information used by nurses to inform practice: An integrative review. Int J Nurs Stud 2008 Jun;45(6):954-970. [Medline: 18677809]

38.   Tagliaventi MR, Mattarelli E. The role of networks of practice, value sharing, and operational proximity in knowledge flows between professional groups. Human Relations 2006 Mar 01;59(3):291-319. [doi: 10.1177/0018726706064175]

39.   LaMendola W, Ballantyne N, Daly E. Practitioner Networks: Professional Learning in the Twenty-First Century. British Journal of Social Work 2009 Feb 24;39(4):710-724. [doi: 10.1093/bjsw/bcp023]

40.   Weston MJ, Estrada NA, Carrington J. Reaping benefits from intellectual capital. Nurs Adm Q 2007;31(1):6-12. [Medline: 17198112]

41.   Orzano AJ, McInerney CR, Scharf D, Tallia AF, Crabtree BF. A knowledge management model: Implications for enhancing quality in health care. J. Am. Soc. Inf. Sci 2008 Feb 01;59(3):489-505. [doi: 10.1002/asi.20763]

42.   Newell S, Edelman L, Scarbrough H, Swan J, Bresnen M. 'Best practice' development and transfer in the NHS: the importance of process as well as product knowledge. Health Serv Manage Res 2003 Feb;16(1):1-12. [doi: 10.1258/095148403762539095] [Medline: 12626022]

43.   Wickramasinghe N. Fostering knowledge assets in healthcare with the KMI model. IJMED 2007;4(1):52-95. [doi: 10.1504/IJMED.2007.011455]

44.   Newell S, Swan J. Trust and inter-organizational networking. Human Relations 2000;53(10):1287-1328.

45.   Ayers LL. Nursing practice as knowledge work within a clinical microsystem. Worcester: University of Massachusetts; 2008. URL: http://escholarship.umassmed.edu/cgi/viewcontent.cgi?article=1008&context=gsn_diss&sei-redir=1#search=%22Nursing%20work%20as%20knowledge%20work%20within%20clinical%20microsystem%20LaFave%22 [WebCite Cache ID 6cYlWHSxD]

46.   Ryan J. A History of the Internet and the Digital future. London: Reaktion Books; 2010.

47.   Chu LF, Young C, Zamora A, Kurup V, Macario A. Anesthesia 2.0: internet-based information resources and Web 2.0 applications in anesthesia education. Curr Opin Anaesthesiol 2010 Apr;23(2):218-227. [doi: 10.1097/ACO.0b013e328337339c] [Medline: 20090518]

48.   Grier D, Campbell M. A social history of Bitnet and Listserv, 1985-1991. IEEE Annals Hist. Comput 2000 Apr;22(2):32-41. [doi: 10.1109/85.841135]

49.   Sophia van Zyl A. The impact of Social Networking 2.0 on organisations. The Electronic Library 2009 Nov 13;27(6):906-918. [doi: 10.1108/02640470911004020]

50.   Grajales F. University of British Columbia. 2012. Social Media: a comprehensive knowledge synthesis and case studies of applications in Medicine and Health(care) URL: http://www.collectionscanada.gc.ca/obj/thesescanada/vol2/BVAU/TC-BVAU-42806.pdf [accessed 2015-10-26] [WebCite Cache ID 6cYmIQiHF]

51.   Paton C, Luquel L. IMIA Social Media Working Group. URL: https://imiasocialmedia.wordpress.com/about/ [accessed 2015-02-14] [WebCite Cache ID 6WKFGBTs0]

52.   Kemp S. Slideshare. 2016. Digital in 2016 URL: http://www.slideshare.net/wearesocialsg/digital-in-2016 [accessed 2016-03-17] [WebCite Cache ID 6g5T5xGnI]

53.   Young C. Community management that works: how to build and sustain a thriving online health community. J Med Internet Res 2013;15(6):e119 [FREE Full text] [doi: 10.2196/jmir.2501] [Medline: 23759312]

54.   Kane G, Alavi M, Labianca G, Borgatti S. What's different about social media neworks? A framework and research agenda. MIS Quarterly 2014;38(1):275-304.

55. Oliver YG. Forrester. 2008. Top enterprise Web 2.0 predictions for 2008 URL: http://www.forrester.com/Top+Enterprise+Web+20+Predictions+For+2008/fulltext/-/E-RES43882?docid=43882&src=44213pdf [WebCite Cache ID 6cYmdqLiI]

56. Whittemore R, Knafl K. The integrative review: updated methodology. J Adv Nurs 2005 Dec;52(5):546-553. [doi: 10.1111/j.1365-2648.2005.03621.x] [Medline: 16268861]

57. Elliott D. Reviewing the literature. In: Schneider Z, Whitehead D, Elliott D, LoBondo-Wood G, Haber J, editors. Nursing & midwifery research: methods and critical appraisal for evidence-based practice, 3rd edition. Sydney: Mosby; 2007:46-61.

58. CASP International. Critical Appraisal Tools. 2013. Qualitative Research Checklist URL: http://www.caspinternational.org/mod_product/uploads/CASP%20Qualitative%20Research%20Checklist%2031.05.13.pdf [WebCite Cache ID 6WWP50105]

59. Krippendorff K. Content Analysis: an introduction to its methodology. In: An introduction to its methodology, 2nd edition. Thousand Oaks, California: Sage; 2004.

60. Zhang Y, Wildemuth BM. Qualitative analysis of content. In: Wildemuth BM, editor. Applications of social research methods to questions in information and library science. Westport: Libraries Unlimited; 2009:308-319.

61. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]

62. SIGN. 2008. Scottish Intercollegiate Guidelines Network URL: http://www.sign.ac.uk/index.html [accessed 2015-10-30] [WebCite Cache ID 6cevHavhL]

63. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009 Jul 21;339:b2535 [FREE Full text] [doi: 10.1136/bmj.b2535]

64. De Wever B, Schellens T, Valcke M, Van Keer H. Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. Computers & Education 2006 Jan;46(1):6-28. [doi: 10.1016/j.compedu.2005.04.005]

65. Morken T, Bull N, Moen BE. The activity on a Norwegian Occupational Health mailing list 1997-2006. Occup Med (Lond) 2009 Jan;59(1):56-58 [FREE Full text] [doi: 10.1093/occmed/kqn143] [Medline: 19001070]

66. Rodriguez-Recio FJ, Sendra-Portero F. Analysis of the Spanish-speaking mailing list RADIOLOGIA. Eur J Radiol 2007 Jul;63(1):136-143. [doi: 10.1016/j.ejrad.2007.01.031] [Medline: 17344009]

67. Berman Y. Discussion groups on the Internet as sources of information: the case of social work. ASLib Proceedings 1996;48(2):31-36.

68. Abrahamson K, Fox R, Anderson J. What nurses are talking about: content and community within a nursing online forum. Stud Health Technol Inform 2013;183:350-355. [doi: 10.3233/978-1-61499-203-5-350]

69. Whitaker S, Cox AR, Alexander AM. Internet networking for pharmacists: an evaluation of a mailing list for UK pharmacists. International Journal of Pharmacy Practice 2003 Mar 01;11(1):25-32. [doi: 10.1211/002235702784]

70. Long S, de JD, Ziviani J, Jones A. Paediatricots: Utilisation of an Australian list serve to support occupational therapists working with children. Aust Occup Ther J 2009 Feb;56(1):63-71. [doi: 10.1111/j.1440-1630.2007.00716.x] [Medline: 20854490]

71. Macdonald L, MacPherson DW, Gushulak BD. Online communication as a potential travel medicine research tool: analysis of messages posted on the TravelMed listserv. J Travel Med 2009;16(1):7-12. [doi: 10.1111/j.1708-8305.2008.00255.x] [Medline: 19192121]

72. Bowers L. Constructing international professional identity: what psychiatric nurses talk about on the Internet. Int J Nurs Stud 1997 Jun;34(3):208-212. [Medline: 9219053]

73. Smith C. A Longitudinal Study of the Culture of MEDLIB-L. Journal of Hospital Librarianship 2004 Mar 24;4(1):29-43. [doi: 10.1300/J186v04n01_03]

74. Murty SA, Gilmore K, Richards KA, Altilio T. Using a LISTSERV™ to develop a community of practice in end-of-life, hospice, and palliative care social work. J Soc Work End Life Palliat Care 2012;8(1):77-101. [doi: 10.1080/15524256.2011.652857] [Medline: 22424385]

75. Brynolf A, Johansson S, Appelgren E, Lynoe N, Edstedt Bonamy A. Virtual colleagues, virtually colleagues—physicians' use of Twitter: a population-based observational study. BMJ Open 2013 Jul 24;3(7):e002988. [doi: 10.1136/bmjopen-2013-002988]

76. Cervantez Thompson Teresa L. You've got mail: rehabilitation nurses on the RehabNurse-L LISTSERV. Rehabil Nurs 2002;27(4):146-151. [Medline: 12116527]

77. Mishori R, Levy B, Donvan B. Twitter use at a family medicine conference: analyzing #STFM13. Fam Med 2014 Sep;46(8):608-614 [FREE Full text] [Medline: 25163039]

78. Matta R, Doiron C, Leveridge MJ. The dramatic increase in social media in urology. J Urol 2014 Aug;192(2):494-498. [doi: 10.1016/j.juro.2014.02.043] [Medline: 24576656]

79. Foong Deborah P S, McGrouther DA. An Internet-based discussion forum as a useful resource for the discussion of clinical cases and an educational tool. Indian J Plast Surg 2010 Jul;43(2):195-197 [FREE Full text] [doi: 10.4103/0970-0358.73448] [Medline: 21217980]

80. Chaudhry A, Glodé LM, Gillman M, Miller RS. Trends in twitter use by physicians at the american society of clinical oncology annual meeting, 2010 and 2011. J Oncol Pract 2012 May;8(3):173-178 [FREE Full text] [doi: 10.1200/JOP.2011.000483] [Medline: 22942812]

81.   Murray P. Subject: talk.to/reflect : reflection and practice in nurses' computer-mediated communications. London: Open University; 2001. URL: http://www.nursing-informatics.com/PhD/PMurray.pdfJanuary [accessed 2016-06-10] [WebCite Cache ID 6i9xmtq6t]

82.   Brooks F, Scott P. Exploring knowledge work and leadership in online midwifery communication. J Adv Nurs 2006 Aug;55(4):510-520. [doi: 10.1111/j.1365-2648.2006.03937.x] [Medline: 16866846]

83.   Burg MA, Adorno G, Hidalgo J. An analysis of Social Work Oncology Network Listserv Postings on the Commission of Cancer's distress screening guidelines. J Psychosoc Oncol 2012;30(6):636-651. [doi: 10.1080/07347332.2012.721484] [Medline: 23101548]

84.   Moorley CR, Chinn T. Nursing and Twitter: creating an online community using hashtags. Collegian 2014;21(2):103-109. [Medline: 25109208]

85.   Ferguson C, Inglis SC, Newton PJ, Cripps Peter J S, MacDonald PS, Davidson PM. Social media: a tool to spread information: a case study analysis of twitter conversation at the Cardiac Society of Australia & New Zealand 61st annual scientific meeting 2013. Collegian 2014;21(2):89-93. [Medline: 25109206]

86.   Anderson G, Gleeson S, Rissel C, Wen LM, Bedford K. Twitter tweets and twaddle: twittering at AHPA. National Health Promotion Conference. Health Promot J Austr 2014 Aug;25(2):143-146. [doi: 10.1071/HE13112] [Medline: 25200470]

87.   Desai T, Shariff A, Shariff A, Kats M, Fang X, Christiano C, et al. Tweeting the meeting: an in-depth analysis of Twitter activity at Kidney Week 2011. PloS one 2012;7(7):e40253.

88.   Hew KF, Hara N. Knowledge sharing in online environments: A qualitative case study. J Am Soc Inf Sci 2007 Dec;58(14):2310-2324. [doi: 10.1002/asi.20698]

89.   McKendrick DR, Cumming G, Lee A. Increased use of Twitter at a medical conference: a report and a review of the educational opportunities. J Med Internet Res 2012;14(6):e176 [FREE Full text] [doi: 10.2196/jmir.2144] [Medline: 23232765]

90.   Neill A, Cronin JJ, Brannigan D, O'Sullivan R, Cadogan M. The impact of social media on a major international emergency medicine conference. Emerg Med J 2014 May;31(5):401-404. [doi: 10.1136/emermed-2012-202039] [Medline: 23423992]

91.   Reutzel TJ, Patel R. Medication management problems reported by subscribers to a school nurse listserv. J Sch Nurs 2001 Jun;17(3):131-139. [Medline: 11885443]

92.   Murray PJ. Nurses' computer-mediated communications on NURSENET. A case study. Comput Nurs 1996;14(4):227-234. [Medline: 8718843]

93.   Hajar Z, Clauson KA, Jacobs RJ. Analysis of pharmacists' use of Twitter. Am J Health Syst Pharm 2014 Apr 15;71(8):615-619. [doi: 10.2146/ajhp130307] [Medline: 24688034]

94.   Schoch NA, Shooshan SE. Communication on a listserv for health information professionals: uses and users of MEDLIB-L. Bull Med Libr Assoc 1997 Jan;85(1):23-32 [FREE Full text] [Medline: 9028568]

95.   Watson D A R. Ozbug: an email mailing list for physicians that works. Intern Med J 2003 Nov;33(11):532-534. [Medline: 14656258]

96.   Cook-Craig PG, Sabah Y. The Role of Virtual Communities of Practice in Supporting Collaborative Learning among Social Workers. British Journal of Social Work 2009 May 12;39(4):725-739. [doi: 10.1093/bjsw/bcp048]

97.   Hughes B, Joshi I, Lemonde H, Wareham J. Junior physician's use of Web 2.0 for information seeking and medical education: a qualitative study. Int J Med Inform 2009 Oct;78(10):645-655. [doi: 10.1016/j.ijmedinf.2009.04.008] [Medline: 19501017]

98.   Kim C, Kang BS, Choi HJ, Lee YJ, Kang GH, Choi WJ, et al. Nationwide online social networking for cardiovascular care in Korea using Facebook. J Am Med Inform Assoc 2014;21(1):17-22 [FREE Full text] [doi: 10.1136/amiajnl-2012-001465] [Medline: 23703828]

99.   Fuoco M, Leveridge MJ. Early adopters or laggards? Attitudes toward and use of social media among urologists. BJU Int 2015 Mar;115(3):491-497. [doi: 10.1111/bju.12855] [Medline: 24981247]

100.   Frisch N, Atherton P, Borycki E, Mickelson G, Cordeiro J, Novak LH, et al. Growing a professional network to over 3000 members in less than 4 years: evaluation of InspireNet, British Columbia's virtual nursing health services research network. J Med Internet Res 2014;16(2):e49 [FREE Full text] [doi: 10.2196/jmir.3018] [Medline: 24566806]

101.   Deen SR, Withers A, Hellerstein DJ. Mental health practitioners' use and attitudes regarding the Internet and social media. J Psychiatr Pract 2013 Nov;19(6):454-463. [doi: 10.1097/01.pra.0000438184.74359.88] [Medline: 24241499]

102.   Klee D, Covey C, Zhong L. Social media beliefs and usage among family medicine residents and practicing family physicians. Fam Med 2015 Mar;47(3):222-226 [FREE Full text] [Medline: 25853534]

103.   Tunnecliff J, Ilic D, Morgan P, Keating J, Gaida JE, Clearihan L, et al. The acceptability among health researchers and clinicians of social media to translate research evidence to clinical practice: mixed-methods survey and interview study. J Med Internet Res 2015;17(5):e119 [FREE Full text] [Medline: 25995192]

104.   Loeb S, Bayne CE, Frey C, Davies BJ, Averch TD, Woo HH, American Urological Association Social Media Work Group. Use of social media in urology: data from the American Urological Association (AUA). BJU Int 2014 Jun;113(6):993-998 [FREE Full text] [doi: 10.1111/bju.12586] [Medline: 24274744]

105.   Hoffmann T, Desha L, Verrall K. Evaluating an online occupational therapy community of practice and its role in supporting occupational therapy practice. Aust Occup Ther J 2011 Oct;58(5):337-345. [doi: 10.1111/j.1440-1630.2011.00954.x] [Medline: 21957918]

106. Kukreja P, Heck SA, Riggins J. Use of social media by pharmacy preceptors. Am J Pharm Educ 2011 Nov 10;75(9):176 [FREE Full text] [doi: 10.5688/ajpe759176] [Medline: 22171104]

107. Usher WT. Australian health professionals' social media (Web 2.0) adoption trends: early 21st century health care delivery and practice promotion. Aust J Prim Health 2012;18(1):31-41. [doi: 10.1071/PY10084] [Medline: 22394660]

108. Stewart S, Sidebotham M, Davis D. International networking: connecting midwives through social media. Int Nurs Rev 2012 Sep;59(3):431-434. [doi: 10.1111/j.1466-7657.2012.00990.x] [Medline: 22897197]

109. Apostolakis I, Koulierakis G, Berler A, Chryssanthou A, Varlamis I. Use of social media by healthcare professionals in Greece: an exploratory study. Int J Electron Healthc 2012;7(2):105-124. [doi: 10.1504/IJEH.2012.049873] [Medline: 23079026]

110. von MM, Ohno-Machado L. Reviewing social media use by clinicians. J Am Med Inform Assoc 2012;19(5):777-781 [FREE Full text] [doi: 10.1136/amiajnl-2012-000990] [Medline: 22759618]

111. Lawson C, Cowling C. Social media: The next frontier for professional development in radiography. Radiography 2015 May;21(2):e74-e80. [doi: 10.1016/j.radi.2014.11.006]

112. Benetoli A, Chen TF, Aslani P. The use of social media in pharmacy practice and education. Res Social Adm Pharm 2015 Feb;11(1):1-46. [doi: 10.1016/j.sapharm.2014.04.002] [Medline: 24814268]

113. Grindrod K, Forgione A, Tsuyuki RT, Gavura S, Giustini D. Pharmacy 2.0: a scoping review of social media use in pharmacy. Res Social Adm Pharm 2014;10(1):256-270. [doi: 10.1016/j.sapharm.2013.05.004] [Medline: 23810653]

114. Roberts MJ, Perera M, Lawrentschuk N, Romanic D, Papa N, Bolton D. Globalization of continuing professional development by journal clubs via microblogging: a systematic review. J Med Internet Res 2015;17(4):e103 [FREE Full text] [doi: 10.2196/jmir.4194] [Medline: 25908092]

115. Hamm MP, Chisholm A, Shulhan J, Milne A, Scott SD, Klassen TP, et al. Social media use by health care professionals and trainees: a scoping review. Acad Med 2013 Sep;88(9):1376-1383. [doi: 10.1097/ACM.0b013e31829eb91c] [Medline: 23887004]

116. Moorhead S, Hazlett D, Harrison L, Carroll J, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]

117. Kaplan AM, Haenlein M. Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons 2010 Jan;53(1):59-68. [doi: 10.1016/j.bushor.2009.09.003]

118. Rolls KD, Hansen M, Jackson D, Elliott D. Analysis of the social network development of a virtual community for Australian intensive care professionals. Comput Inform Nurs 2014 Nov;32(11):536-544. [doi: 10.1097/CIN.0000000000000104] [Medline: 25310223]

119. Lulic I, Kovic I. Analysis of emergency physicians' Twitter accounts. Emerg Med J 2013 May;30(5):371-376. [doi: 10.1136/emermed-2012-201132] [Medline: 22634832]

120. McGowan BS, Wasko M, Vartabedian BS, Miller RS, Freiherr DD, Abdolrasulnia M. Understanding the factors that influence the adoption and meaningful use of social media by physicians to share medical information. J Med Internet Res 2012;14(5):e117 [FREE Full text] [doi: 10.2196/jmir.2138] [Medline: 23006336]

121. Lau Adela SM. Hospital-based nurses' perceptions of the adoption of Web 2.0 tools for knowledge sharing, learning, social interaction and the production of collective intelligence. J Med Internet Res 2011;13(4):e92 [FREE Full text] [doi: 10.2196/jmir.1398] [Medline: 22079851]

122. Archambault PM, Bilodeau A, Gagnon M, Aubin K, Lavoie A, Lapointe J, et al. Health care professionals' beliefs about using wiki-based reminders to promote best practices in trauma care. J Med Internet Res 2012;14(2):e49 [FREE Full text] [doi: 10.2196/jmir.1983] [Medline: 22515985]

123. Thompson TL, Penprase B. RehabNurse-L: an analysis of the rehabilitation nursing LISTSERV experience. Rehabil Nurs 2004;29(2):56-61. [Medline: 15052747]

124. Hew KF, Hara N. An online listserv for nurse practitioners: a viable venue for continuous nursing professional development? Nurse Educ Today 2008 May;28(4):450-457. [doi: 10.1016/j.nedt.2007.07.009] [Medline: 17881096]

125. Stewart SA, Abidi Syed Sibte Raza. Applying social network analysis to understand the knowledge sharing behaviour of practitioners in a clinical online discussion forum. J Med Internet Res 2012;14(6):e170 [FREE Full text] [doi: 10.2196/jmir.1982] [Medline: 23211783]

126. Rolls K, Kowal D, Elliott D, Burrell AR. Building a statewide knowledge network for clinicians in intensive care units: knowledge brokering and the NSW Intensive Care Coordination and Monitoring Unit (ICCMU). Aust Crit Care 2008 Feb;21(1):29-37. [doi: 10.1016/j.aucc.2007.10.003] [Medline: 18226542]

127. Dieleman C, Duncan E. Investigating the purpose of an online discussion group for health professionals: a case example from forensic occupational therapy. BMC health services research 2013;13(1):253. [doi: 10.1186/1472-6963-13-253]

128. Valaitis RK, Akhtar-Danesh N, Brooks F, Binks S, Semogas D. Online communities of practice as a communication resource for community health nurses working with homeless persons. J Adv Nurs 2011 Jun;67(6):1273-1284. [doi: 10.1111/j.1365-2648.2010.05582.x] [Medline: 21306424]

129. Roberts C, Fox N. General practitioners and the Internet: modelling a 'virtual community'. Fam Pract 1998 Jun;15(3):211-215 [FREE Full text] [Medline: 9694177]

130. Thomas RE, James SD. Informal communications networking among health professionals: a study of gp-uk. Health Informatics Journal 1999 Jun 01;5(2):74-81. [doi: 10.1177/146045829900500204]

131. Mishori R, Singh LO, Levy B, Newport C. Mapping physician Twitter networks: describing how they work as a first step in understanding connectivity, information flow, and message diffusion. J Med Internet Res 2014;16(4):e107 [FREE Full text] [doi: 10.2196/jmir.3006] [Medline: 24733146]

132. Probst G, Borzillo S. Why communities of practice succeed and why they fail. European Management Journal 2008 Oct;26(5):335-347. [doi: 10.1016/j.emj.2008.05.003]

133. Lai H, Chen TT. Knowledge sharing in interest online communities: A comparison of posters and lurkers. Computers in Human Behavior 2014 Jun;35:295-306. [doi: 10.1016/j.chb.2014.02.004]

134. Dann S. First Monday. 2010. Twitter content classification URL: http://firstmonday.org/ojs/index.php/fm/article/viewArticle/2745/2681, [WebCite Cache ID 6fLjm90Hm]

135. Awad NI, Cocchio C. Use of Twitter at a major national pharmacy conference. Am J Health Syst Pharm 2015 Jan 1;72(1):65-69. [doi: 10.2146/ajhp140175] [Medline: 25511841]

136. Hawkins CM, Duszak R, Rawson JV. Social media in radiology: early trends in Twitter microblogging at radiology's largest international meeting. J Am Coll Radiol 2014 Apr;11(4):387-390. [doi: 10.1016/j.jacr.2013.07.015] [Medline: 24139963]

137. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]

138. Scott P, Brooks F, Quick K, Macintyre M, Rospopa C. Assisted Electronic Communication in Nursing. Hertfordshire; 2004. URL: http://kmi.open.ac.uk/publications/pdf/kmi-04-9.pdf [WebCite Cache ID 6i8MZvt7f]

139. Brooks F, Rospopa C, Scott P. Midwifery on the net: new communication technology. British Journal of Midwifery 2004;12(2):107-110.

140. Dahlander L, Frederiksen L. The Core and Cosmopolitans: A Relational View of Innovation in User Communities. Organization Science 2012 Aug;23(4):988-1007. [doi: 10.1287/orsc.1110.0673]

141. Ikioda F, Kendall S, Brooks F, De Liddo A, Buckingham Shum S. Factors That Influence Healthcare Professionals' Online Interaction in a Virtual Community of Practice. SN 2013;2(4):174-184. [doi: 10.4236/sn.2013.24017]

142. Irvine-Smith S. A Series of Encounters: The Information Behaviour of Participants in a Subject-Based Electronic Discussion List. J Info Know Mgmt 2010 Sep;9(3):183-201. [doi: 10.1142/S0219649210002619]

143. Chen G, Chiu MM. Online discussion processes: Effects of earlier messages' evaluations, knowledge content, social cues and personal information on later messages. Computers & Education 2008 Apr;50(3):678-692. [doi: 10.1016/j.compedu.2006.07.007]

144. Wasko M, Faraj S. Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. MIS Quarterly 2005;29(1):35-57.

145. Lin MJ, Hung S, Chen C. Fostering the determinants of knowledge sharing in professional virtual communities. Computers in Human Behavior 2009 Jul;25(4):929-939. [doi: 10.1016/j.chb.2009.03.008]

146. Chiu C, Hsu M, Wang ET. Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. Decision Support Systems 2006 Dec;42(3):1872-1888. [doi: 10.1016/j.dss.2006.04.001]

147. Hsu M, Ju TL, Yen C, Chang C. Knowledge sharing behavior in virtual communities: The relationship between trust, self-efficacy, and outcome expectations. International Journal of Human-Computer Studies 2007 Feb;65(2):153-169. [doi: 10.1016/j.ijhcs.2006.09.003]

148. Cho H, Chen M, Chung S. Testing an integrative theoretical model of knowledge-sharing behavior in the context of Wikipedia. J. Am. Soc. Inf. Sci 2010;61(6):1198-1212. [doi: 10.1002/asi.21316]

149. Bock G, Zmud R, Kim Y, Lee J. Behavioural intention formation in knowledge sharingxaming the roles of extrinsic motivators, social-psychological forces and organizational climate. MIS Quarterly 2005;29(1):87-111.

150. Hew K, Hara N. Empirical study of motivators and barriers of teacher online knowledge sharing. Educational Technology, Research and Development 2007;55:573-595.

151. Rangachari P, Rissing P, Wagner P, Rethemeyer K, Mani C, Bystrom C, et al. A baseline study of communication networks related to evidence-based infection prevention practices in an intensive care unit. Qual Manag Health Care 2010;19(4):330-348. [doi: 10.1097/QMH.0b013e3181fa36a4] [Medline: 20924254]

152. Dopson S, FitzGerald L, Ferlie E, Gabbay J, Locock L. No magic targets! Changing clinical practice to become more evidence based. Health Care Manage Rev 2002;27(3):35-47. [Medline: 12146782]

153. Hara N. Information technology support for communities of practice: How public defenders learn about winning and losing in court. J Am Soc Inf Sci 2006 Jan 01;58(1):76-87. [doi: 10.1002/asi.20445]

154. Bullen M. Participation and Critical Thinking in Online University Distance Education. Journal of Distance Education 1998;13(2):1-32.

155. Cadogan M, Nickson C. Life in the Fast Lane 13 June. 2014. FOAM URL: http://lifeinthefastlane.com/foam/ [WebCite Cache ID 6i8NAt33n]

156. Li X, Gray K, Chang S, Elliott K, Barnett S. A conceptual model for analysing informal learning in online social networks for health professionals. Stud Health Technol Inform 2014;204:80-85. [Medline: 25087531]

Abbreviations

CoP: community of practice
HCP: health care professionals
SM: social media
SNS: social networking sites
TRIP: translating research into practice
VC: virtual community

XSL•FO
RenderX

Original Paper

Trial Promoter: A Web-Based Tool for Boosting the Promotion of Clinical Research Through Social Media

Katja Reuter[1*], PhD; Francis Ukpolo[1], MBA; Edward Ward[1], BS; Melissa L Wilson[2], PhD, MPH; Praveen Angyan[1*], MS

[1]Southern California Clinical and Translational Science Institute, Keck School of Medicine of USC, University of Southern California, Los Angeles, CA, United States

[2]Departments of Preventive Medicine & Obstetrics and Gynecology, Keck School of Medicine of USC, University of Southern California, Los Angeles, CA, United States

[*]these authors contributed equally

Corresponding Author:
Katja Reuter, PhD
Southern California Clinical and Translational Science Institute
Keck School of Medicine of USC
University of Southern California
CSC 200 Bldg, 2nd floor
2250 Alcazar Street
Los Angeles, CA, 90033
United States
Phone: 1 3234422046
Fax: 1 3234422082
Email: Katja.Reuter@usc.edu

Abstract

Background: Scarce information about clinical research, in particular clinical trials, is among the top reasons why potential participants do not take part in clinical studies. Without volunteers, on the other hand, clinical research and the development of novel approaches to preventing, diagnosing, and treating disease are impossible. Promising digital options such as social media have the potential to work alongside traditional methods to boost the promotion of clinical research. However, investigators and research institutions are challenged to leverage these innovations while saving time and resources.

Objective: To develop and test the efficiency of a Web-based tool that automates the generation and distribution of user-friendly social media messages about clinical trials.

Methods: Trial Promoter is developed in Ruby on Rails, HTML, cascading style sheet (CSS), and JavaScript. In order to test the tool and the correctness of the generated messages, clinical trials (n=46) were randomized into social media messages and distributed via the microblogging social media platform Twitter and the social network Facebook. The percent correct was calculated to determine the probability with which Trial Promoter generates accurate messages.

Results: During a 10-week testing phase, Trial Promoter automatically generated and published 525 user-friendly social media messages on Twitter and Facebook. On average, Trial Promoter correctly used the message templates and substituted the message parameters (text, URLs, and disease hashtags) 97.7% of the time (1563/1600).

Conclusions: Trial Promoter may serve as a promising tool to render clinical trial promotion more efficient while requiring limited resources. It supports the distribution of any research or other types of content. The Trial Promoter code and installation instructions are freely available online.

KEYWORDS

Introduction

Scarce information about clinical research, in particular clinical trials, is among the top reasons why potential participants do not take part in clinical trials. Clinical trials are vital for the development of novel approaches to advancing medicine, but without volunteers this type of research is impossible. In 2012, the Institute of Medicine recognized the seriousness of the clinical trial participation problem [1] and released a report that identified numerous barriers, including the lack of awareness among patients and physicians that clinical trials are available. New solutions are needed that increase clinical trial awareness and build rapport among patients, physicians, and caregivers with the aim to boost clinical trial engagement and recruitment rates.

We have developed and tested Trial Promoter, a Web-based tool that automatically generates and distributes social media messages about clinical trials. New digital options such as social media (ie, social networks) have the potential to work alongside traditional methods to boost the promotion of clinical research. Social media describe websites and Web-based applications that enable users to create and share content or to participate in social networking. With millions of users, social media serve as a promising solution to improve the public awareness of clinical trials and to support research participant recruitment efforts. In fact, the use of the Internet as a top source for clinical research information has increased significantly (46% in 2013), whereas the use of mass media has declined (newspaper, radio, television; 39% in 2013) [2]. Between 30% and 40% of the public reports that they have used social media to gather medical information to learn about clinical research, with the social network Facebook topping the list [3-6]. This trend is not limited to young adults; half of people aged 50 years and older and more than a third of people aged 65 years and older frequent social networking sites such as Twitter and Facebook [7-9]. These data suggest that patients, caregivers, and disease advocates can be found, informed, and engaged digitally. We have built Trial Promoter to leverage this digital trend and to support research institutions that seek to respond to the evolving way in which patients, physicians, caregivers, and advocacy groups search for, create, and use health information online.

Trial Promoter builds on preliminary work where we tested an automated approach to generate and publish messages about research-related content on Twitter. Our work indicated that a machine algorithm helps research teams and institutions to increase the output and reach of information about research on social media while reducing the burden of developing and distributing hundreds of messages [10].

Here we present the tool and preliminary data, suggesting that Trial Promoter may aid in distributing clinical trial information more broadly while requiring limited resources. The tool serves four functions: first, it imports information from specific databases or data files. Second, it generates user-friendly social media messages based on preapproved message templates. Third, it schedules and distributes these messages through the social media platforms Twitter and Facebook. Fourth, it tracks the success of the messages and displays their engagement and conversion metrics data. The source code and installation instructions are freely available online [11]. In order to test the tool, we conducted a 10-week trial. Trial Promoter randomized 46 active and recruiting clinical trials into social media messages and distributed them via Twitter and Facebook. We assessed the correctness of the test messages and calculated the probability with which Trial Promoter generated accurate messages.

Methods

Trial Promoter Setup

System Requirements

Trial Promoter is built using Ruby on Rails (version 4.2.1), HTML, cascading style sheet (CSS), and JavaScript. We have installed Trial Promoter on Ubuntu Linux 14.04 LTS (long-term support) and use Phusion Passenger, a scalable Web server that hosts Trial Promoter. Trial Promoter further uses PostgreSQL 9.4.5 database systems deployed on Ubuntu 14.04 LTS. Administrator privileges for setting up Cron jobs are required in order to set up nightly data extractions that secure logs, collate metrics, and distribute messages.

Information Import

Trial Promoter has the capability to import information from different databases and data files through either a representational state transfer (REST) application programming interface (API) or a comma-separated values (CSV) file. Figure 1 depicts the Trial Promoter setup and data flow. Figure 2 represents a screenshot of the local Trial Promoter interface that shows imported clinical trial information and disease keywords that were included in the test messages. Table 1 lists the information our local Trial Promoter installation imported for testing purposes, for example, clinical trial information from our institutional Clinical Studies Directory that utilizes data from ClinicalTrials.gov provide by the National Library of Medicine [12], message templates designed for Twitter and Facebook, and information on disease hashtags. Disease hashtags are disease keywords preceded by a pound sign (eg, #Diabetes, #BreastCancer). They are used by members of disease communities on social media sites such as Twitter to identify and discover messages on a specific topic [13,14].

Table 1. Data sources and types of data imported for testing purposes by our local installation of Trial Promoter.

Imported content	Data source/format	Data types
Active, recruiting clinical trials	Clinical Studies Directory/REST API[a]	Clinical trial title
		Name of principal investigator
		Clinical trial landing page URL
Disease hashtags	Symplur [11]/CSV[b] file	Disease keywords
Message templates	N/A[c]/CSV file	Text message templates designed for Twitter and Facebook

[a]REST API: representational state transfer application programming interface.

[b]CSV: comma-separated values.

[c]N/A: not applicable.
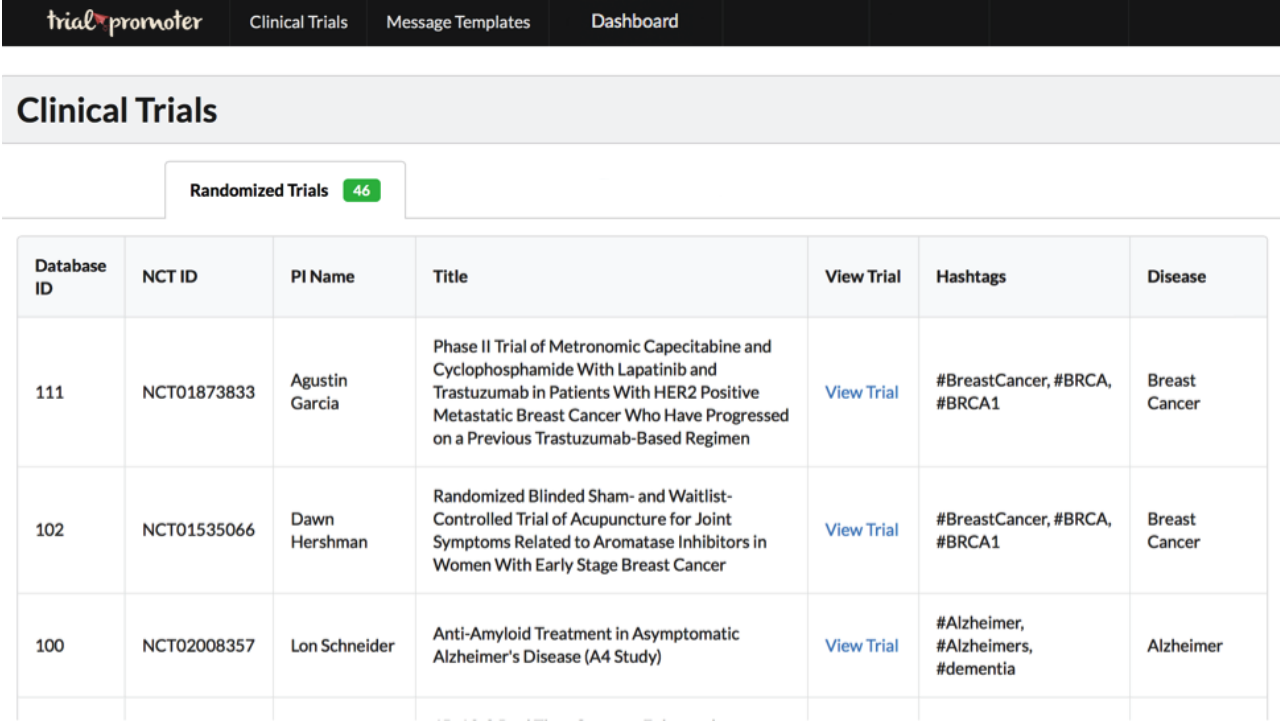
Figure 1. Trial Promoter (TP) setup and data flow. The elements in blue represent functional TP modules. CSV: comma-separated values; REST API: representational state transfer application programming interface.

Figure 2. Local Trial Promoter interface that shows imported clinical trial information and disease keywords that were included in the test messages.

Message Generation and Distribution

Trial Promoter Messenger develops social media messages tailored to Twitter and Facebook. To achieve that, Trial Promoter uses preapproved message templates that we developed with input from communication experts at Keck Medicine of the University of Southern California (USC). During testing, we used 154 parameterized message templates. Figure 3 represents a screenshot of the local Trial Promoter interface that shows parameterized message templates for Twitter and Facebook, which were used during testing. The parameterization supports the generation of large sets of messages from a limited set of clinical trials and message templates. The words in italics the following example messages represent parameters that Trial Promoter added into the message templates to create the final social media messages.

Example of Twitter message template: "New #ClinicalTrial @*KeckMedUSC* on *#disease* is looking for participants. Please help us spread the word. Thx. *URL* "

Example of Facebook message template: "Your help is appreciated: New clinical trial at Keck Medicine of USC on *#disease* is looking for participants. Through these types of clinical studies researchers can better understand how to diagnose, treat and prevent diseases. Please share this *URL*. Thanks!"

Trial Promoter matches a randomly chosen clinical trial with a randomly chosen message template using the standard Ruby library to generate random numbers [15]. The random numbers in the library are implemented as a modified Mersenne Twister with a period of $2^{19937}-1$ [16]. Trial Promoter shuffles all clinical trials and then randomly chooses a message template for each trial [17], ensuring that all clinical trials are only distributed once during a given time period.

Trial Promoter then substitutes the parameters in the message templates and includes several weblinks into the message templates to create the final messages: first, a tagged and shortened URL that links the social media message to the respective clinical trial landing page; second, a primary and if applicable a secondary disease hashtag (eg, #LungCancer, #SleepApnea); and third, for Twitter messages the official Keck Medicine of USC Twitter account (@KeckMedUSC). Table 2 lists the characteristics of the messages that Trial Promoter generated automatically during the testing phase.

Table 2. Characteristics of test messages that Trial Promoter generated automatically for distribution on Twitter and Facebook.

Characteristic	Twitter	Facebook
Maximum message length limitation	Limitation to 140 characters[a]	N/A[b]
Parameter: URL	22 characters for non-https URLs 23 characters for https URLs	Links can be of any length. However, in order to simplify URL sharing and present clean URLs to the Facebook page visitor, Trial Promoter uses Bit.ly shortened URLs on Facebook posts as well.
Parameter: hashtags (disease keyword)	Yes (primary and if length permits secondary hashtag)	Yes (primary and secondary hashtags)
Parameter: link to official Keck Medicine of USC[c] Twitter account (@KeckMedUSC)	Yes	N/A[b]

[a]Note: Media attachments such as photos, videos, and polls are not counted toward 140 characters.

[b]N/A: not applicable.
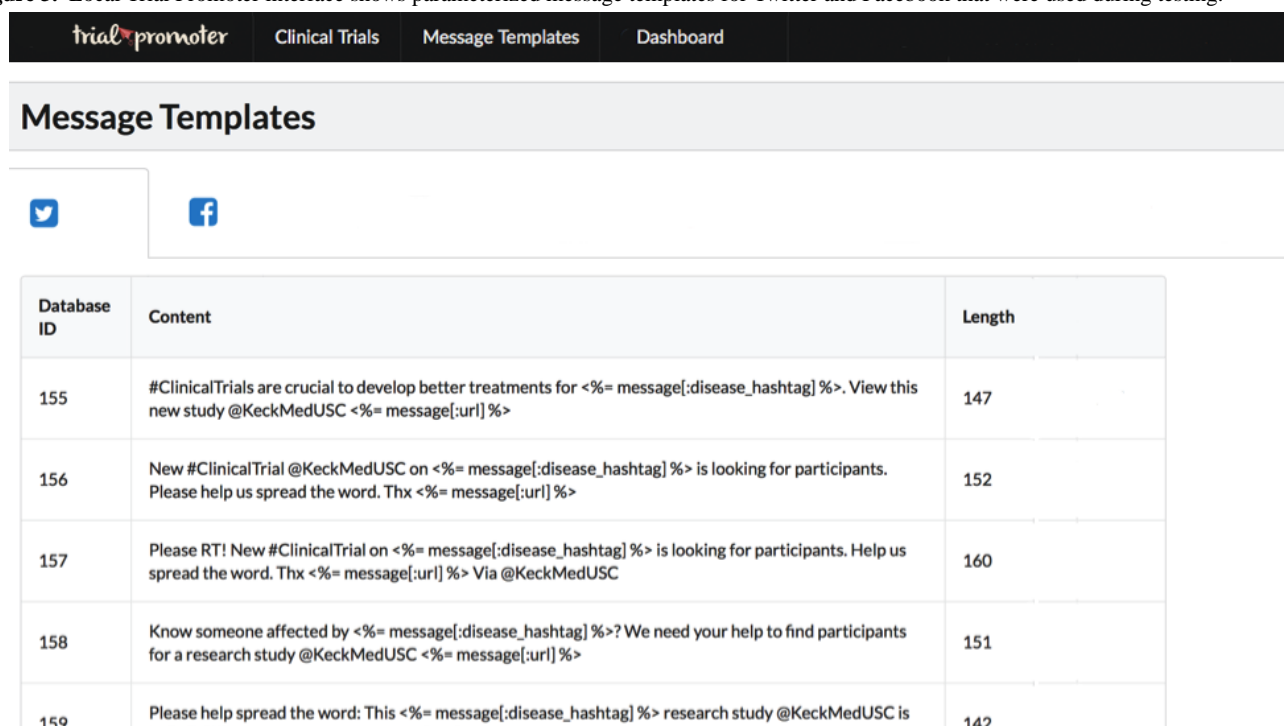
[c]USC: University of Southern California.

Furthermore, Trial Promoter tags the URL that links to the clinical trial landing page with Urchin Traffic Monitor (UTM) parameters in order to track the link engagement (or clicks) on social media and referral traffic to the clinical trial landing page [18]. During testing, Trial Promoter used the REST API provided by the Bit.ly link-shortening service to generate the shortened URL [19]. Bit.ly preserves the UTM parameters by mapping identical links with different UTM parameters to unique URLs.

For Twitter, the automated inclusion of disease hashtags that vary in length depending on the disease term used may result in messages that are longer than 140 characters (eg, #HIV vs #PancreaticCancer). If the generated message was greater than 140 characters, Trial Promoter discarded the message and selected an alternative message template until it either generated a valid message or it ran out of message templates to choose from. In the latter case, Trial Promoter notified the study team of the error in the administrative dashboard.

Trial Promoter Messenger then schedules and distributes the test messages through selected Twitter and Facebook accounts (eg, USC Clinical Trials) using the social media content management Web application Buffer [20]. Each social media account set up in Buffer has a unique profile identifier (ID) assigned to it. Buffer provides a REST API call that allows for programmatic scheduling of messages directly in Buffer. With a single call, the Buffer REST API [21] sends a single message to multiple social media channels. The REST API provides options for scheduling messages and including images. Once messages are sent to Buffer, the tool interface allows users to edit, reschedule, reorder, and delete messages.

Figure 3. Local Trial Promoter interface shows parameterized message templates for Twitter and Facebook that were used during testing.

Automated Tracking of Message Success

Trial Promoter Collector tracks key performance indicators (KPIs) associated with the success of the messages, that is, social media engagement and landing page conversion data. Table 3 lists the engagement and conversion KPIs tracked by Trial Promoter. Trial Promoter Collector extracts these data on a nightly basis from 4 existing systems: Buffer, Twitter, Facebook, and Google Analytics [22]. The tool connects to Buffer via a REST API in order to extract the social media and link engagement data for each social media message sent via Buffer. To collect the data from Twitter and Facebook, our team uploaded reports taken from each social media platform to Trial Promoter. Trial Promoter processed these raw logs and imported them into its dashboard on a nightly basis. The landing page conversion data are accessible through Google Analytics. To extract the data, Trial Promoter required uploads of the Google Analytics data report into Trial Promoter. Trial Promoter then processed and incorporated the data into its dashboard.

Table 3. Engagement and conversion key performance indicators tracked by Trial Promoter.

Metric categories	Measures on Twitter	Measures on Facebook
Volume of messages served	Impressions	Impressions
Social media engagement	Retweets	Shares
	Replies	Comments
	Likes	Likes
Link engagement	Clicks from social media message to clinical trial landing page on Clinical Studies Directory	Clicks from social media message to clinical trial landing page on Clinical Studies Directory
Landing page engagement	Sessions	Sessions
	Time spent on page	Time spent on page
	Pageviews per visit	Pageviews per visit
Contact engagement	Contact form usage on individual clinical trial information page	Contact form usage on individual clinical trial information page

Finally, *Trial Promoter Dashboard* displays the KPI data, providing daily updates. Figure 4 represents a screenshot of the local Trial Promoter interface that shows KPI data for each Twitter message. During testing, Trial Promoter matched up data from the raw data logs for each social media platform to the data from Buffer using a unique social message (or update) ID. Each social media channel generates a unique ID for every message that is sent out via its platform. Using this unique ID allowed us to match up entries in the raw data logs to a specific message. Additionally, the Trial Promoter Dashboard serves as a control panel to add, edit, and delete clinical trials, message templates, and social media messages and images.

Trial Promoter Evaluation

Test Trial Design

During the 10-week test trial, Trial Promoter randomized clinical trials (n=46) into social media messages using preapproved message templates. The tool generated, scheduled, and published 2 messages on each platform (Twitter and Facebook) per day, and 3 messages per platform every other day.

Correctness Analysis

The correctness with which Trial Promoter generated messages during the 10-week testing phase was evaluated using 4 indicators: (1) the correct usage of the message template, (2) the text of the message itself (ie, number of text errors in the message), (3) the inclusion of the correct URL, and (4) the inclusion of the correct disease hashtags. The individual indicators were averaged to obtain the overall percent correct. The correct usage of the message template was measured through random sampling 25/525 messages (5%) and manually comparing the message template with the social media that was generated by Trial Promoter. The correctness of the included URL was evaluated using a script written in Ruby on Rails [8]. The script examined the URL by first expanding the shortened Bit.ly URL to a complete URL. The URL was then compared with the landing page of the clinical trial being promoted using regex expressions (ignoring any query strings in the URL) to ensure that they were identical. The inclusion of the correct disease hashtags (eg, #Stroke, #LungCancer) was manually reviewed in all 525 messages upon scheduling.

Figure 4. Local Trial Promoter interface shows key performance indicator data for each message. This example shows Twitter messages.

Results

Trial Promoter Evaluation

During the 10-week testing phase, Trial Promoter successfully generated and distributed 175 messages via Twitter and 350 messages via Facebook, a total of 525 messages. Figure 5 shows examples of messages that Trial Promoter generated automatically for distribution on social media.

Correctness Analysis

The analysis of the test messages revealed that in a random sample of 25/525 messages the Trial Promoter algorithm used the correct message template without any text errors 100% of the time. However, we discovered a recurring issue in 13/525 messages (2.5%) where the algorithm used a question mark instead of an apostrophe (eg, "We're looking for participants"). Furthermore, the analysis of the URLs showed that 525/525 messages (100%) included the correct URL. Finally, the disease hashtag analysis revealed that 24/525 distributed messages (4.6%) lacked the disease hashtag. Trial Promoter had not substituted the parameter "#disease" in the message template with the disease of the clinical trial in all cases. On average, Trial Promoter correctly used the message templates and substituted the message parameters (text, URLs, and disease hashtags) 97.7% of the time (1563/1600).

Availability

The Trial Promoter software code is available under the MIT license [22,23]. Software code versions for technical and nontechnical users are accessible through the Trial Promoter website and hosted on GitHub [10]. To simplify the installation process for nontechnical users, we have written a script that deploys an instance of Trial Promoter to the public hosting service Heroku—within less than 30 minutes and without requiring technical knowledge of server setup and system administration. A fee of US $14 per month is required for hosting Trial Promoter on Heroku. Using the provided code, nontechnical users will be able to do the following: add clinical trial information for promotion—one trial at a time; import message templates—one at a time; automatically integrate with the social media management tool Buffer that automates the distribution of the messages; and create a dashboard that imports metrics from Buffer via a REST API, and from Twitter and Facebook via a CSV file. For technical users, detailed instructions for hosting Trial Promoter on an Ubuntu 14.04 LTS machine are also available on GitHub. Using the provided code, technical users will be able to do the following: import clinical

trial information via a REST API—many trials at once; import message templates via a REST API—multiple templates at once; automatically integrate with Buffer that automates the distribution of the messages; and create a dashboard that automatically imports metrics from Buffer via a REST API, and that imports metrics from Twitter and Facebook via a CSV file. Finally, the full code used in the experiment described here is accessible through the Trial Promoter website and viewable on GitHub [10].

Figure 5. Examples of messages that Trial Promoter generated and published automatically.
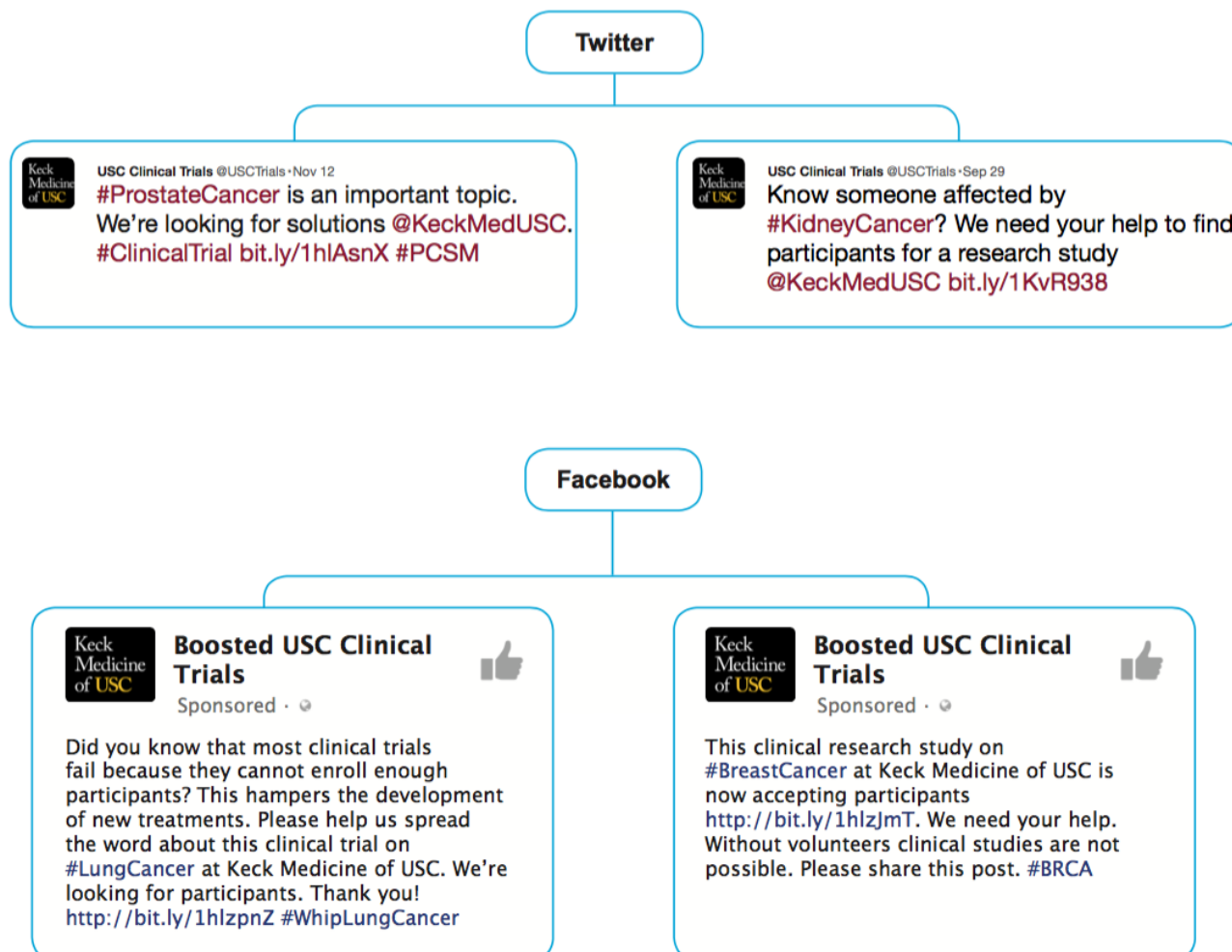
Discussion

Automated Clinical Trial Promotion

We have developed and tested Trial Promoter, a Web-based tool that automatically generates and distributes user-friendly social media messages about clinical trials. We chose the social media platforms Facebook and Twitter because members of disease communities frequently use them [2,24,25].

Trial Promoter managed to generate and distribute social media messages with a high level of correctness. We were able to improve the Trial Promoter algorithm with regard to the technical issues we encountered during testing. First, in 5% of the messages the algorithm had not substituted the generic disease hashtag parameter (#disease) with the corresponding disease hashtag taken from the clinical trial. We found that the substitution algorithm was both case and white-space sensitive (ie, # Disease, # disease, and #Disease). As a result, only the parameter #disease without a space and with lowercase "d" was replaced with the actual disease term that was associated with a clinical trial. We have modified the Trial Promoter algorithm, rendering the replacement of disease hashtags case and white-space insensitive, thereby resolving the issue for the future.

In light of millions of messages divulged by social media users on Twitter and Facebook every day, tailoring those messages to a specific audience is critical to cut through the noise. Hashtags provide a useful tool to target messages to specific topics and disease communities. The simple # symbol, known as a hashtag, for example, #leukemia, #rheum, is included in each message to indicate a topic, conversation, or event on Twitter [14] that the message relates to. Our ongoing work is focused on assessing the effectiveness of Trial Promoter for the promotion of clinical trial messages through a new Twitter account without followers. Following a Twitter user means to subscribe to a person's feed, that is, stream of messages.

Second, in 2.5% of the messages the algorithm introduced a question mark instead of an apostrophe owing to encoding issues. This error was introduced while copying the message templates from a Google Docs file [24] into the code of our local Trial Promoter installation that we used during the testing

phase. We consider this a rare technical issue specific to Google Docs where an apostrophe is encoded differently than text within a code editor. However, if necessary this type of technical issue can be addressed and corrected through the administrative dashboard of Trial Promoter.

The data thus far indicate that Trial Promoter serves as a promising tool to support clinical trial promotion via social media. More specifically, Trial Promoter is designed to facilitate two phases of the clinical trial recruitment process. Figure 6 illustrates these phases: the promotion (or advertisement) phase and the engagement phase. By publishing messages about clinical trials on social media, Trial Promoter supports the awareness-building phase of the clinical trial recruitment process. Several studies indicate that awareness changes attitudes toward clinical trials, enrollment, and the benefits of participation. More than 80% of patients were either unaware or unsure that participation in a clinical trial was an option at the time of diagnosis, and 75% of these patients said they would have been willing to enroll had they known it was possible [25,26]. These data indicate that improving the distribution of clinical trial information at limited cost may benefit clinical trial recruitment efforts.

Trial Promoter further generates opportunity for social media engagement by potential study participants, disease advocates, and others because social media is designed to facilitate interaction and conversation, for example, sharing, liking, following, and replying. The link engagement in the message that Trial Promoter generates and distributes is essential to triage visitors to the clinical trial landing page where they can find more information about the trial and potentially contact the study team using a compliant contact form.

Figure 6. Trial Promoter is designed to facilitate two phases of the clinical trial recruitment process: the promotion (advertisement) and engagement phases.
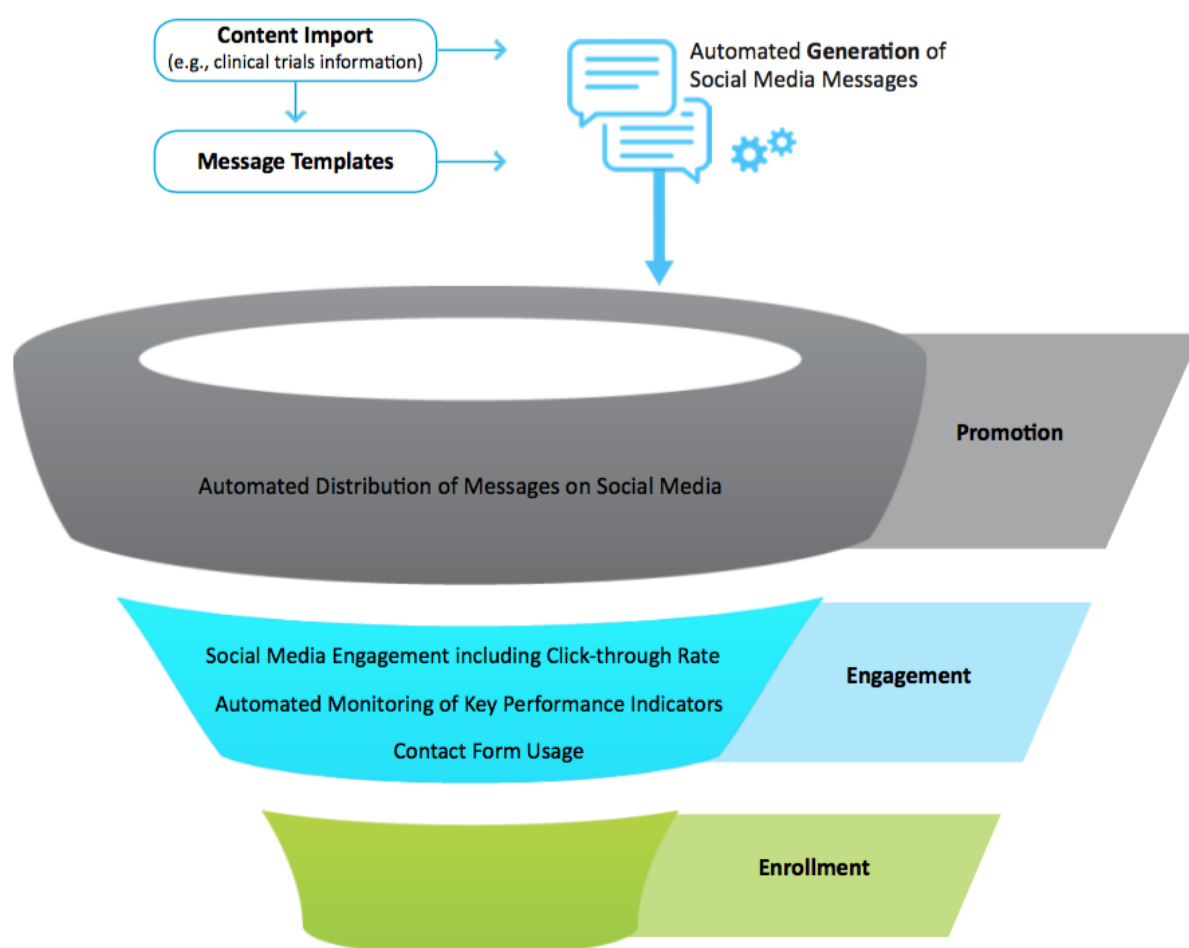
Saving Time and Cost

We wondered whether Trial Promoter could increase fiscal efficiencies of the clinical trial promotion. This is especially relevant in light of the high operational costs associated with clinical trials [4,27]. Nearly 30% of the time dedicated to clinical trials is spent on promotion, patient recruitment, and enrollment [28]. Despite this substantial amount of time and cost, more than 30% of all clinical trials fail to meet their enrollment targets, and more than 10% never enroll a single patient [29].

To test the theory as to whether Trial Promoter makes the promotion of clinical trials on social media more efficient fiscally, we compared the labor and cost of Trial Promoter with the labor and cost of a social media manager. During the 10-week testing phase, Trial Promoter automatically generated 525 user-friendly messages. The generation of a social media

message takes an average of 5-15 minutes including selecting the content, writing the message, adding an optional image or video, and posting the message (based on internal analysis). According to jobs and recruiting site Glassdoor, the national average annual salary for a social media manager in 2015 was US $52,000 [30]. A social media manager would have required 43-131 hours of labor equivalent of US $1800-$5400 labor cost to generate the number of messages Trial Promoter has generated in 10 weeks (525). We argue that Trial Promoter helps to decrease the time and cost of labor required to generate and distribute information about clinical trials through social media. Although there were costs associated with the development of Trial Promoter (ie, personnel cost, server cost), those are minimal in comparison to the fact that the tool will be a free resource to the research community. It can be used repeatedly at no additional cost to investigators and research institutions. The tool may also help research institutions and investigators to gain efficiencies by streamlining and improving the clinical trials infrastructure and process so that those investigating new research questions could quickly draw on resources already in place instead of reinventing the wheel for each trial—in this case the broader dissemination of clinical trial information via social media.

Saving time and cost would not—in part—be possible without automation. However, the use of automated postings in digital marketing and social media is a controversial topic. We believe automation to serve patients, disease advocates, and research institutions alike when used appropriately and as long as the content is relevant and of value to the audience. Trial Promoter, for example, allows clinical study teams and others in charge of promoting clinical trials to design and write the posts at times when it is more convenient for them. Trial Promoter thereby not only saves them time, the tool also schedules and disseminates the social media posts at times when it is more convenient for the audience.

Related Work

A number of studies have discussed the automation of the production of news and information. Automation offers new possibilities for creating content at scale, more quickly than a human could. So-called "bots," that is, automated accounts on digital and social media (eg, Twitter, Facebook, Reddit, and Wikipedia) that distribute news and information, have been observed and studied in a variety of contexts: in social networks and human communication decisions [31,32], social shaping [33], content pollution [34], social metric gaming [35], ranking manipulation [36], infiltration [37], political astroturfing [38], recommendation [39], scholarship dissemination [40], activism or advocacy [41], and journalism [42]. Lokot and Diakopoulos concluded that news bots might enable innovation, such as niche and local news [42]. Different definitions have been introduced to describe these bots as "automated social actors"—software designed to act similarly to how humans might act in social spaces [31], as "software agents that interact on social networking services" [33], and as "automatic or semi-automatic computer programs that mimic humans and/or human behavior" [43]. However, future research will need to investigate how the public perceives news and information bots, whether they recognize bots as automated information services, if they are

skeptical of content shared by a bot, and whether bots are ultimately effective in achieving the bottom line, for example, increase clinical study recruitment or foster the accessibility of public health information. Some small-scale work suggests that "while the software-generated content is perceived as descriptive and boring, it is also considered to be objective although not necessarily discernible from content written by journalists" [44]. Another study found that a Twitter bot sharing public health information was perceived as "credible," "attractive," and "competent," suggesting that such "bots could be gainfully employed by [organizations] if properly harnessed" [45]. The authors are not aware of similar research that has tested the feasibility and effectiveness of automated postings of clinical research information. Yet the preliminary data in other fields look promising.

Limitations

Limitations of the Study

The test trial we present here was focused on assessing the probability with which Trial Promoter generates and distributes correct messages about clinical trials. Future studies will be required to systematically assess the efficacy of Trial Promoter (or other algorithms) beyond fiscal efficiencies, determining the ability of machine-generated clinical trial information to foster the awareness of and engagement among target audiences such as patients, disease advocates, and physicians. Furthermore, it will be interesting to explore to what extent social media engagement with machine-generated content translates into increased clinical trial recruitment and enrollment rates.

Limitations of Trial Promoter

The current version of Trial Promoter does not automatically include images into the social media messages. Images, however, have been shown to be an important aspect of social media messages to increase engagement and information uptake [46]. The tool also does not automatically include mentions of influencers in the messages, that is, names of Twitter users with a lot of followers and reach—an important technique to increase the exposure of messages among target audiences. Additionally, Trial Promoter does not yet take into account awareness months when scheduling messages, for example, October is Breast Cancer Awareness Month. To increase the reach of the distributed information, Trial Promoter could increase the promotion of disease-related messages during awareness months.

Conclusions

In summary, we present Trial Promoter and preliminary data indicating that the tool reliably automates the generation and distribution of correct clinical trial messages via social media. The Trial Promoter software code is freely available online. Although our local installation and pilot project focuses on clinical trials, Trial Promoter has the capability to support the generation and distribution of any type of content. Other examples of content include research news stories, peer-reviewed articles, and information about research experts and their expertise.

We hypothesize that machine-generated content helps research institutions and investigators to distribute clinical trial

information more broadly and effectively. However, further studies around machine-generated content on social media will help to understand its role in facilitating patient engagement, increasing clinical trial awareness, and improving study recruitment and retention rates.

Conflicts of Interest

None declared.

References

1. Public Engagement and Clinical Trials: New Models and Disruptive Technologies: Workshop Summary. Washington, DC: National Academies Press (US); 2012.
2. The Center for Information & Study on Clinical Research Participation. 2013. Report on clinical trial information seekers. Perceptions and insights study URL: https://www.ciscrp.org/ [accessed 2016-05-26] [WebCite Cache ID 6hn2TkTmy]
3. The Center for Information & Study on Clinical Research Participation. 2015. Report on the decision to participate. Perceptions and insights study URL: https://www.ciscrp.org/ [accessed 2016-05-26] [WebCite Cache ID 6hn2acue8]
4. Taylor H, Leitman R. Misconceptions and lack of awareness greatly reduce recruitment for cancer clinical trials. Harris Interactive: Health Care News 2001;1(3):1-3.
5. Fox S, Jones S. Pew Research Center. The Social Life of Health Information URL: http://www.pewinternet.org/files/old-media/Files/Reports/2009/PIP_Health_2009.pdf [accessed 2016-02-18] [WebCite Cache ID 6fNNA4ZdY]
6. Facebook. URL: https://facebook.com/ [accessed 2016-02-17] [WebCite Cache ID 6fNFRSsU0]
7. Smith A. Pew Research Center. Older adults and technology use URL: http://www.pewinternet.org/2014/04/03/older-adults-and-technology-use/ [accessed 2016-02-17] [WebCite Cache ID 6fNFn66n4]
8. Fox S, Duggan M. Pew Research Center.: Pew Research Center; 2013. Health Online 2013 URL: http://www.pewinternet.org/2013/01/15/health-online-2013/ [accessed 2016-02-17] [WebCite Cache ID 6fNNRtbVO]
9. Twitter. URL: https://twitter.com/ [accessed 2016-02-17] [WebCite Cache ID 6fNFYrrhm]
10. Reuter K, Chatterjee A, Daigre J, Voytek B. Spreading research and engaging disease communities: One automated tweet at a time. 2014 Presented at: Proceedings of the Joint Summits on Translational Science; April; 2014; San Francisco, United States p. 7-11.
11. Trial Promoter. 2016. URL: http://trialpromoter.org [accessed 2016-02-17] [WebCite Cache ID 6fMwI58NC]
12. Clinicaltrials.gov. U.S. National Institutes of Health URL: http://clinicaltrials.gov/ [accessed 2016-02-17] [WebCite Cache ID 6fNHGi4Zu]
13. Symplur. Healthcare Hashtag Project URL: http://www.symplur.com/healthcare-hashtags/ [accessed 2016-02-17] [WebCite Cache ID 6fNG11EjX]
14. Harmel M, Young K. e-Patients in twitter hashtag communities. J Participat Med 2013;5:e22.
15. Ruby documentation project. URL: http://ruby-doc.org/core-2.2.0/Random.html [accessed 2016-02-16] [WebCite Cache ID 6fNHNVE0H]
16. Matsumoto M, Nishimura T. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. ACM Trans Model Comput Simul 1998;8(1):3-30. [doi: 10.1145/272991.272995]
17. Ruby documentation project. Shuffle URL: http://ruby-doc.org/core-2.2.0/Array.html#method-i-shuffle [accessed 2016-02-17] [WebCite Cache ID 6fNHpzEdI]
18. Google. Add parameters to URLs to identify the campaigns that refer traffic URL: https://support.google.com/analytics/answer/1033863?hl=en [accessed 2016-02-17] [WebCite Cache ID 6fNHyqvir]
19. https://bitly.com/. Bitly URL: https://bitly.com/ [accessed 2016-02-17] [WebCite Cache ID 6fNHuGMkh]
20. https://buffer.com. Buffer URL: https://buffer.com [accessed 2016-02-17] [WebCite Cache ID 6fNI55AJe]
21. https://buffer.com/developers/api/updates. Buffer Rest API URL: https://buffer.com/developers/api/updates [accessed 2016-02-17] [WebCite Cache ID 6fNI81yL5]
22. Google. Google Analytics URL: https://www.google.com/analytics/ [accessed 2016-02-17] [WebCite Cache ID 6fNIDa8lO]
23. Creative Commons. MIT License URL: http://choosealicense.com/licenses/mit/ [accessed 2016-02-17] [WebCite Cache ID 6fNH2emCR]

24. Utengen A. Symplur. 2012 Dec 10. The rise of patient communities on Twitter ? Twitter Visualized URL: http://www.symplur.com/shorts/the-rise-of-patient-communities-on-twitter-visualized/#!prettyPhoto/0/ [accessed 2016-02-17] [WebCite Cache ID 6fNGr1HmB]

25. Utengen A. Symplur. 2012 Dec 17. The Dynamics of a Twitter Patient Community – Network Centrality Analysis URL: http://www.symplur.com/shorts/dynamics-twitter-patient-community-network-centrality-analysis/ [accessed 2016-02-17] [WebCite Cache ID 6fNGh78Hu]

26. Google. Google Docs URL: https://www.google.com/docs/ [accessed 2016-02-17] [WebCite Cache ID 6fNGXzA6u]

27. Temkar P. Clinical operations generation next… The age of technology and outsourcing. Perspect Clin Res 2015;6(4):175-178 [FREE Full text] [doi: 10.4103/2229-3485.167098] [Medline: 26623386]

28. Garrett A, Brown Stafford P. Using Real-time Data to Drive Better Decisions, Faster. Therapeutic Innovation & Regulatory Science 2011;45(4):495-502. [doi: 10.1177/009286151104500410]

29. Tufts University. 2013. Impact Report, Tufts Center for the Study of Drug Development URL: http://csdd.tufts.edu/files/uploads/02_-_jan_15,_2013_-_recruitment-retention.pdf [accessed 2016-05-26] [WebCite Cache ID 6hn4grtTS]

30. Glassdoor. Social Media Manager Salaries URL: https://www.glassdoor.com/Salaries/social-media-manager-salary-SRCH_KO0,20.htm [accessed 2016-02-17] [WebCite Cache ID 6fNGOy3Na]

31. Norah A, Yoo D, Mcdonald D. Dissecting a social botnet. In: CSCW. 2015 Presented at: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work and Social Computing; March 14 - 18, 2015; Vancouver, BC, Canada p. 839-851.

32. Tavares G, Faisal A. Scaling-laws of human broadcast communication enable distinction between human, corporate and robot Twitter users. PLoS One 2013 Jul;8(7):e65774 [FREE Full text] [doi: 10.1371/journal.pone.0065774] [Medline: 23843945]

33. Hwang T, Pearce I, Nanis I. Socialbots: Voices from the fronts. Interactions 2012;19(2):38-45.

34. Kyumin L, Eoff B, Caverlee J. Seven months with the devils: A long-term study of content polluters on Twitter. In: Association for the Advancement of Artificial Intelligence Digital Library. Menlo Park, California: The AAAI Press; 2011 Presented at: Proceedings of the 5th AAAI International Conference on Weblogs and Social media; July 17 – 21, 2011; Barcelona, Catalonia, Spain.

35. Cook D, Waugh B, Abdipanah M, Hashemi O, Rahman S. Twitter deception and influence: Issues of identity, slacktivism, and puppetry. Journal of Information Warfare 2014;13(1):58-71.

36. Eni M, Metaxas P. From obscurity to prominence in minutes: Political speech and real-time search. Raleigh, NC: US; 2010 Presented at: Proceedings of the WebSci10: Extending the Frontiers of Society On-Line; April 26-27, 2010; Raleigh, North Carolina.

37. Yazan B, Muslukhov I, Beznosov K, Ripeanu M. The socialbot network: when bots socialize for fame and money. 2011 Presented at: 27th Annual Computer Security Applications Conference; December 05 - 09, 2011; Orlando, FL, USA p. 93-102.

38. Ratkiewicz J, Conover M, Meiss M, Gonçalves B, Flammini A, Menczer F. Detecting and tracking political abuse in social media. In: Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media. 2011 Presented at: 5th International AAAI Conference on Weblogs and Social Media; July 17-21, 2011; Barcelona, Spain p. 297-304.

39. Maria A, Deplano M, Schifanella R, Ruffo G. People Are Strange When You're a Stranger: Impact and Influence of Bots on Social Networks. In: Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media. Palo Alto, California: The AAAI Press; 2012 Presented at: Sixth International AAAI Conference on Weblogs and Social Media; June 4–7, 2012; Dublin, Ireland p. 10-17.

40. Haustein S, Bowman T, Holmberg K, Tsou A, Sugimoto C, Larivière V. Tweets as impact indicators: Examining the implications of automated "bot" accounts on Twitter. J Assn Inf Sci Tec 2015 May 05;67(1):232-238. [doi: 10.1002/asi.23456]

41. Wilkie A, Michael M, Plummer-Fernandez M. Speculative method and Twitter: Bots, energy and three conceptual characters. Sociol Rev 2014 Jul 14;63(1):79-101. [doi: 10.1111/1467-954X.12168]

42. Lokot T, Diakopoulos N. News Bots. Automating news and information dissemination on Twitter. Digital Journalism 2015 Sep 15:1-18. [doi: 10.1080/21670811.2015.1081822]

43. Wagner C, Mitter S, Koerrner C, Strohmaier M. When social bots attack: Modeling susceptibility of users in online social networks. New York: ACM; 2012 Presented at: Proceedings of the 2nd Workshop on Making Sense of Microposts (MSM'2012), held in conjunction with the 21st World Wide Web Conference (WWW'2012); April 16 - 20, 2012; Lyon, France.

44. Clerwall C. Enter the Robot Journalist. Journalism Practice 2014 Feb 25;8(5):519-531. [doi: 10.1080/17512786.2014.883116]

45. Edwards C, Edwards A, Spence P, Shelton A. Is that a bot running the social media feed? Testing the differences in perceptions of communication quality for a human agent and a bot agent on Twitter. Computers in Human Behavior 2014 Apr;33:372-376. [doi: 10.1016/j.chb.2013.08.013]

46. Zarrella D. The science of marketing: When to tweet, what to post, how to blog, and other proven strategies. New York: Wiley; 2013.

Abbreviations

API: application programming interface
CSV: comma-separated values
HIV: human immunodeficiency virus
ID: identifier
KPI: key performance indicator
LTS: long-term support
REST: representational state transfer
USC: University of Southern California
UTM: Urchin Traffic Monitor

XSL•FO
RenderX

Original Paper

Preferred Features of E-Mental Health Programs for Prevention of Major Depression in Male Workers: Results From a Canadian National Survey

JianLi Wang[1,2,3], PhD; Raymond W Lam[4], MD; Kendall Ho[5,6], MD; Mark Attridge[7], PhD; Bonnie M Lashewicz[2], PhD; Scott B Patten[1,2,3], M.D., Ph.D; Alain Marchand[8], PhD; Alice Aiken[9], PhD; Norbert Schmitz[10], PhD; Sarika Gundu[11], MSc; Nitika Rewari[12], MSc; David Hodgins[13], PhD; Andrew Bulloch[1,2,3], PhD; Zul Merali[14,15,16], PhD

[1]Department of Psychiatry, University of Calgary, Calgary, AB, Canada

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

[3]Mathison Centre for Mental Health Research & Education, Hotchkiss Brain Institute, University of Calgary, Calgary, AB, Canada

[4]Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada

[5]Department of Emergency Medicine, University of British Columbia, Vancouver, BC, Canada

[6]Digital Emergency Medicine, University of British Columbia, Vancouver, BC, Canada

[7]Attridge Consulting Inc., Minneapolis, MN, United States

[8]School of Industrial Relations, Public Health Research Institute, University of Montreal, Montreal, QC, Canada

[9]School of Rehabilitation Therapy, Queens University, Kingston, ON, Canada

[10]Department of Psychiatry, McGill University, Montreal, QC, Canada

[11]Workplace Mental Health Program, Canadian Mental Health Association - National, Toronto, ON, Canada

[12]Workplace Mental Health, Mental Health Commission of Canada, Ottawa, ON, Canada

[13]Department of Psychology, University of Calgary, Calgary, AB, Canada

[14]Department of Psychology, University of Ottawa, Ottawa, ON, Canada

[15]Institute of Mental Health Research, The Royal, University of Ottawa, Ottawa, ON, Canada

[16]Department of Psychiatry, University of Ottawa, Ottawa, ON, Canada

Corresponding Author:
JianLi Wang, PhD
Department of Psychiatry
University of Calgary
3280 Hospital Dr. NW
Calgary, AB, T2N4Z6
Canada
Phone: 1 403 2108653
Fax: 1 403 210 9182
Email: jlwang@ucalgary.ca

Abstract

Background: Major depression is a prevalent mental disorder and imposes considerable burden on health and productivity. Men are not immune to major depression, yet they often delay seeking help because of perceived stigma and gender norms. E-mental health programs hold potential for early prevention of major depression. However, we have little knowledge about men's preferences for design features of e-mental health programs.

Objectives: The objective of this study was to (1) estimate and compare the proportions of Internet use for medical information, preferred design features, and likely use of e-mental health programs; (2) examine factors associated with the likely use of e-mental health programs; and (3) understand potential barriers to the use of e-mental health programs among Canadian working men, who were at high risk of a major depressive episode (MDE).

Methods: A cross-sectional survey in 10 Canadian provinces was conducted between March and December 2015. Random digit dialing method was used through household landlines and cell phones to collect data from 511 working men who were at high risk of having an MDE and 330 working men who were at low risk of having an MDE.

Results: High-risk men were more likely to endorse the importance of accessing health resources on the Internet than low-risk men (83.4% vs 75.0%, respectively; *P*=.01). Of the 17 different features assessed, the top three features most likely to be used by high-risk men were: "information about improving sleep hygiene" (61.3%), "practice and exercise to help reduce symptoms of stress and depression" (59.5%), and "having access to quality information and resources about work stress issues" (57.8%). Compared with men at low risk for MDE, men at high risk for MDE were much more likely to consider using almost every one of the different design features. Differences in preferences for the design features by age among men at high risk of MDE were found only for 3 of 17 features. Differences in preferences for design features between English- and French-speaking participants were found only for 4 out of the 17 features. Analysis of qualitative data revealed that privacy issues, perceived stigma, ease of navigation, personal relevance, and lack of personal interaction, time, and knowledge were identified as barriers to the use of e-mental health programs in working men who were at high risk of MDE.

Conclusion: E-mental health programs may be a promising strategy for prevention of depression in working men. Development of e-mental health programs should consider men's preferences and perceived barriers to enhance the acceptability of this approach.

Introduction

Major depression is a prevalent mental disorder. In Canada, the annual prevalence of major depressive disorder was 3.9% based on the DSM-IV criteria [1]; in the United States, the 12-month prevalence was 6.6% [2]. Major depression is disabling and significantly affects workers' health and productivity. Depression alone accounts for 2.5% of the global burden of disease and is among the largest single causes of disability worldwide (8.2% of all years lived with disability globally) [3]. US workers with depression cost an estimated US $44.01 billion per year in lost productivity [4]. Epidemiological studies have consistently found that women are more likely to have major depression than men. However, men are not immune to depression. Canadian national data showed that annually, 2.8% of adult men have a major depressive disorder [1]. One of the severe consequences of having an MDE is suicide. Canadian and US national data showed that 75%-80% of all suicides were male [5,6].

Given its considerable effect on health, productivity, and lives, there is a pressing need for innovation in prevention of major depression in male workers. However, this is a challenging endeavor. In the workplace, risk factors for having an MDE differ for men and women [7-9]. For instance, job strain, family-to-work conflict, and job insecurity appear to be more prominent risk factors for MDE among men than among women [7,8]. Men are less likely than women to seek help and to disclose depressive symptoms and often delay seeking help until symptoms become severe, which compounds these risks. Men tend to socialize, to be emotionally stoic, and exemplify traditional masculine characteristics such as independence, self-reliance, and dominance [10]. Men are concerned about the perceived negative judgments from family, friends, and coworkers if they access treatment for depression. These gender-specific experiences, together with the limited knowledge of effective interventions, call for innovative solutions tailored for men. One of the burgeoning solutions that has attracted considerable interest is e-mental health.

E-mental health is "the use of information and communication technologies to support and improve mental health, including the use of online resources, social media, and mobile phone apps" [11]. E-mental health self-help services enable users to learn more about their mental health conditions and empower them to strengthen their self-management and improve their health. Most e-mental health treatment programs were developed based on the models of cognitive behavioral therapy (CBT) or interpersonal therapy (IPT) that have been proven to be effective in depression and anxiety [12]. A review by Christensen and Petrie showed that by 2013, 62 Web-based mental health interventions and 11 mobile apps had been developed [13]. Lal and Adair found 91 peer-reviewed publications on the application of e-mental health interventions between 2000 and 2010 [14]. Thus far, randomized controlled trials of e-mental health interventions on depression have included individuals with clinical depression or those above a threshold of a depression rating scale [14]. Consistent with public health mandates to reduce the burden of depression, it is important that e-mental health not only addresses needs of those with depression or those who are above a threshold depression rating, but also advances capacity for secondary prevention [15], by identifying high-risk individuals and intervening to prevent early symptoms from progressing into an MDE. Additionally, existing e-mental health programs have not been designed and evaluated using a gender lens. Given the gender norms, the extent to which men accept e-mental health programs is unknown. Effectiveness of the program and its acceptability to users constitute the foundation for scalable and sustainable program implementation. Therefore, as part of our BroMatters study (www.bromatters.ca), we conducted a cross-sectional survey among working men, some of whom were at high risk of MDE, to understand their preference for design features of e-mental health programs.

The objectives of this analysis were to (1) estimate and compare the proportions of Internet use for medical information, preferred design features, and likely use of e-mental health programs, (2) examine factors associated with the likely use of e-mental health programs, and (3) understand potential barriers to the use of

e-mental health programs among working men who were at variable risk levels of MDE.

Methods

Study Design and Recruitment

A cross-sectional survey was conducted between March and December 2015. The target population of the survey included Canadian working men who: (1) were aged 18 years or older, (2) did not have an MDE in the past 12 months, (3) were at high risk of an MDE at the time of interview (a low-risk sample was also obtained for comparison), (4) were working at the time of the survey, and (5) have no language barriers to either English or French. Because of the vast geographic area of Canada, participants were recruited using the random digit dialing method by the Bureau of Professional Interviewers (BIP), located in Montreal, Canada. The BIP has access to household telephone numbers across the country and to a validated mobile phone number database, and its interviewers can conduct interviews in both English and French. The Conjoint Health Research Ethics Review Board of University of Calgary approved the study.

Once a household was reached, the household contact was asked to retrieve or provide contact information (eg, a first name) of the household residents who are men and are currently working. If there was more than one potentially eligible individual in the same household, one of them was randomly selected. Once the prospective participant was fully informed about the objectives and procedures of the study, oral consent was obtained to proceed with the interview. Participants were first administered a risk calculator for MDE to estimate their probability of having an MDE in the future. The definition of high risk is described below. The number of high-risk participants in each age group was proportional to the age distribution of Canadian male working population in 2014, provided by Statistics Canada.

From March to December 2015, 49,500 calls were made. A majority of the calls (47,648, 96.2%) were not valid (not in service, fax or modem, answering machine, language barriers, ineligibility, duplications, refusal before eligibility was assessed). Among 1852 eligible participants, 596 (32.1%) refused to participate after eligibility verification; 842 provided complete data (45.4%). The remaining included incomplete interviews or scheduled call-backs not in study period (22.5%). After removing 1 duplication, 841 participants were included in the analysis, including 511 men who were at high risk of having major depression and 330 who were at low risk of having major depression.

Measurements

A multivariable riskprediction algorithm for major depression in men was administered to estimate the risk (probability) of having an MDE in the next 4 years for each participant [16]. This risk prediction model was designed to be used in individuals who did not have an MDE. Based on the participant's exposure to a set of key risk factors (predictors) in the model, the algorithm can generate the absolute risk or probability of having an MDE in the next 4 years, analogous to the Framingham riskprediction algorithm for coronary heart

disease [17,18]. The risk-prediction algorithm for MDE in men was developed and validated using data of 4737 Canadian men who were aged 18 years or older and did not have an MDE in the past 12 months [16]. The risk-prediction algorithm contains 15 predictors including age, personal and family history of MDE, childhood trauma, ongoing stress and life events, and antidepressant or sleeping pill use in the past month. The predictive power of the risk-prediction algorithm was measured using C statistics (C=0.7953) [16], which is equivalent to the area under the curve when the outcome is binary. The model had excellent calibration with data as indicated by the Hosmer–Lemeshow test and a visual comparison between the predicted and observed risks by decile risk groups [16]. In our study, >6.51% were identified as high risk, which represents the top two decile risk groups in the Canadian male population. Predicted risk lower than 6.51% was defined as low risk. Internet use was assessed using questions from the 2012 Canadian Internet Use Survey conducted by Statistics Canada [19].

Preferred design features of e-mental health program questions were developed by BroMatters team members. Participants were asked questions such as, "We want to hear your opinion about e-mental health programs for dealing with work and stress issues. E-health is defined as …, For the following features, please indicate how likely it is that you would use them . " In all, 17 questions about design features were asked. For each question, participants answered on a 5-point Likert scale ranging from very likely to very unlikely. Open-ended questions were asked about any other features they may want in an e-mental health program, whether the participant and his male coworkers may use an e-mental health program to deal with work stressand what makes it difficult to use an e-mental health program. Administering the questions and instruments to eligible participants took an average of 22 min. Participants who completed the survey received a CAN $20 gift card as a token of appreciation.

Statistical Analysis

The background characteristics and proportions of likely use of design features were estimated and compared between men who were at high or low risk of having an MDE using the chi-square test. Among men who were at high risk of having an MDE, the percentages were also estimated and compared by age groups and by language used in the interview (English vs French) using a chi-square test. All analyses were conducted using the statistical program STATA 14.0 (StataCorp, College Station, TX, USA). Tests were considered statistically significant when p was less than .05. With this level of probability and a sample size of 841, the study had a statistical power of 0.89 to detect a small effect size (Cohen d) of 0.20.

Results

Participants' Characteristics

The demographic and socioeconomic characteristics of the participants are summarized in Table 1. Men who were at high risk of having an MDE had characteristics similar to those at low risk, except that high-risk men were younger and were more likely to have reported work function impairment ($P<.001$).

Table 1. Demographic and socioeconomic characteristics of the participants overall and by risk levels of having major depression.

Variable		Total (N=841), n (%)	High risk (n= 511), n (%)	Low risk (n=330), n (%)	P
Age categories (years)					<.001
	18-29	102 (12.1)	73 (14.3)	29 (8.8)	
	30-49	444 (52.8)	284 (55.6)	160 (48.5)	
	50-64	254 (30.2)	140 (27.4)	114 (34.5)	
	>65	41 (4.9)	14 (2.7)	27 (8.2)	
Mean age (SD)		44.3 (13.7)	42.0 (12.2)	47.8 (15.0)	
Marital status					.14
	Married or Common-law	647 (77.0)	389 (76.1)	258 (78.4)	
	Divorced or separated/widowed	43 (5.1)	22 (4.3)	21 (4.4)	
	Single	150 (17.9)	100 (19.6)	50 (15.2)	
Personal income ($)					.60
	<30,000	99 (12.2)	60 (12.1)	39 (12.4)	
	30,000 to <60,000	238 (29.4)	152 (30.6)	86 (27.4)	
	60,000 to 80,000	171 (21.1)	98 (19.7)	73 (23.3)	
	>80,000	303 (37.4)	187 (37.6)	116 (36.9)	
Educational levels					.15
	Above high school	65 (7.7)	43 (8.4)	22 (6.7)	
	High school	166 (19.7)	107 (20.9)	59 (17.9)	
	College	265 (31.5)	167 (32.7)	98 (29.7)	
	University or higher	345 (41.0)	194 (38.0)	151 (45.8)	
Employment					.70
	Employee	673 (80.6)	413 (81.5)	260 (79.3)	
	Self-employed	159 (19.0)	92 (18.2)	67 (20.4)	
	Family business no pay	3 (0.4)	2 (0.4)	1 (0.3)	
Job type					.16
	Full time	692 (82.3)	434 (84.9)	258 (78.2)	
	Part time	71 (8.4)	37 (7.2)	34 (10.3)	
	Seasonal	37 (4.4)	18 (3.5)	19 (5.8)	
	Contract	36 (4.3)	19 (3.7)	17 (5.2)	
	Other	5 (0.6)	3 (0.6)	2 (0.6)	
Size of company or work site					.78
	<50	445 (53.6)	276 (54.2)	169 (52.5)	
	50-499	243 (29.2)	149 (29.3)	94 (29.2)	
	>500	143 (17.2)	84 (16.5)	59 (18.3)	
Provinces					.22
	British Columbia	47 (6.3)	33 (7.1)	14 (5.0)	
	Alberta	70 (9.4)	52 (11.2)	18 (6.4)	
	Saskatchewan	30 (4.0)	18 (3.9)	12 (4.3)	
	Manitoba	31 (4.2)	20 (4.3)	11 (3.9)	
	Ontario	293 (39.3)	170 (36.6)	123 (43.4)	
	Quebec	227 (30.5)	141 (30.4)	86 (30.6)	

XSL•FO
RenderX

Variable		Total (N=841), n (%)	High risk (n= 511), n (%)	Low risk (n=330), n (%)	P
	New Brunswick	16 (2.2)	12 (2.6)	4 (1.4)	
	Nova Scotia	17 (2.3)	12 (2.6)	5 (1.8)	
	Newfoundland	10 (1.3)	4 (0.9)	6 (2.1)	
	Prince Edward Island	4 (0.5)	2 (0.4)	2 (0.7)	
Language					.81
	English	613 (72.9)	374 (73.2)	239 (72.4)	
	French	228 (27.1)	137 (26.8)	91 (27.6)	
Work function impairment					<.001
	None	577 (70.4)	314 (62.9)	263 (81.9)	
	Mild	199 (24.3)	146 (29.3)	53 (16.5)	
	Moderate	41 (5)	36 (7.2)	5 (1.6)	
	Severe	3 (0.4)	3 (0.6)	0	

Internet Usage

A majority of the participants reported use of Internet for personal reasons in the past 12 months, with the proportion (95.7%) among high-risk men being slightly higher than that among low-risk men (92.4%) (Table 2). The two groups did not differ in Internet use for searching medical information and in perceived usefulness of the Internet information in making decisions about health. However, high-risk men (83.4%) were more likely to report that it was important to access health resources on the Internet than low-risk men (75.0%).

Table 2. General and health-related Internet usage among men who were at different risk levels of major depression.

Internet use during the past 12 months		High risk, n (%)	Low risk, n (%)	P
Used Internet for personal use		489 (95.7)	305 (92.4)	.04
Hours of Internet each week (h)				
	<5	175 (35.8)	110 (36.1)	.81
	5-9	137 (28.0)	93 (30.5)	
	10-19	113 (23.1)	66 (21.6)	
	20-29	39 (8.0)	23 (7.5)	
	30-39	11 (2.3)	8 (2.6)	
	>40	14 (2.9)	5 (1.6)	
Used Internet on a mobile phone, tablet, or other mobile devices		408 (83.4)	235 (77.1)	.03
Used Internet for medical or health-related information		307 (62.7)	168 (55.1)	.08
How useful Internet helps you in making decisions about your health				
	Not useful	32 (10.7)	14 (8.5)	.90
	Unsure	41 (13.4)	28 (16.7)	
	Useful	231 (75.9)	125 (74.8)	
How important is it for you to be able to access health resources on the Internet				
	Not important	31 (10.1)	15 (8.9)	.01
	Unsure	20 (6.5)	27 (16.1)	
	Important	256 (83.4)	126 (75.0)	

Preferred Design Features

Participants rated their level of interest in possible use of 17 different features that can be incorporated into the design of e-mental health programs. We ranked the preferred design features of e-mental health program in descending order (see Table 3). The top three features that were identified by high-risk men as things they would likely to use were: "information about improving sleep hygiene," "practice and exercise to help reduce symptoms of stress and depression," and "having access to quality information and resources about work stress issues." The proportions of individuals endorsing the selected design features were significantly higher in the high-risk group than those in the low-risk group, except for "information about improving sleep hygiene" (Table 3).

Table 3. Proportions of preferred design features of e-mental health program in men who were at different risk levels of major depression.

Features	High risk, n (%)	Low risk, n (%)	P
Likely use of …			
Information about improving sleep hygiene	313 (61.3)	181 (54.9)	.07
Practice and exercise to reduce stress	303 (59.5)	127 (38.8)	<.001
Quality information about work stress	295 (57.8)	146 (44.2)	<.001
Setting personal goals and track them	277 (54.6)	128 (38.8)	<.001
Watching videos online on how to deal with work and stress issues	272 (53.3)	125 (37.9)	<.001
Being able to access a program via mobile phone or as an app.	264 (52.0)	127 (38.6)	<.001
Direct referral to health professional to deal with work and stress issues in person	263 (51.7)	109 (33.0)	<.001
Being able to ask questions and receive answers from mental health professional	253 (49.7)	120 (36.5)	<.001
Self-help interactive program that provides info about work problems	244 (47.8)	96 (29.4)	<.001
Being able to chart and track your mood	217 (42.6)	93 (28.5)	<.001
A risk calculator predicting future risk of having major depression	215 (42.6)	94 (28.8)	<.001
Access by phone to a trained coach to help with work stress	211 (41.3)	97 (29.6)	.001
Information about anger management	211 (41.5)	88 (26.8)	<.001
Receiving printed materials	167 (32.7)	78 (23.7)	.005
Information delivered in game format	156 (30.7)	62 (18.8)	<.001
Online peer connection	140 (27.5)	61 (18.5)	.003
Online chat room	129 (25.3)	58 (17.6)	.009

Preferred Design Features by Age and Language Among High-Risk Men

We estimated and compared the proportions of preferred design features by age groups and languages used in the interviews in men who were at high risk of MDE. The data showed that, compared to older participants, younger participants preferred access to a program through a smartphone or mobile app and that the information be delivered in game format (Table 4).

Middle-aged men preferred receiving printed materials. The preferences for other design features did not vary by age.

English-speaking participants were more likely to use "practice & exercise to reduce stress," "access a program via smartphone or an app," and "being able to ask questions and receive answers from mental health professionals" than French-speaking men; French-speaking men were more likely to use "being able to chart and track your mood" than English-speaking participants (Table 5).

Table 4. Proportions of preferred design features of e-mental health program in men who were at high-risk levels of major depression by age groups.

Features	18-29 y, n (%)	30-49 y, n (%)	50-64 y, n (%)	65+ y, n (%)	P
Likely use of …					
Being able to access a program via smartphone or as an app.	46 (63.0)	156 (55.3)	58 (41.7)	4 (28.6)	.003
Receiving printed materials through mail	15 (20.6)	89 (31.3)	61 (43.6)	2 (14.3)	.002
Information to be delivered in game format	36 (49.3)	84 (29.7)	33 (23.74)	3 (21.4)	.001

Table 5. Proportions of preferred design features of e-mental health program in men who were at high risk levels of major depression, by language used.

Features	English, n (%)	French, n (%)	P
Likely use of …			
Practice and exercise to reduce stress	235 (63.2)	68 (49.6)	.006
Being able to access a program via smartphone or as an app	206 (55.4)	58 (42.7)	.01
Being able to ask questions and receive answers from mental health professionals	197 (52.8)	56 (41.2)	.02
Being able to chart and track your mood	147 (39.5)	70 (51.1)	.02

A majority of participants considered our survey questions, about preferred design features, to be comprehensive and did not have other features to add. For the open-ended questions, some participants suggested that, in addition to the design features encompassed in the survey, other valuable features may be: easy to use (eg, "online information in a format that is simple to use."), confidentiality (eg, "Privacy, somehow to ensure privacy"), credibility (eg, "having access to reliable information that's important to me"), and direct link to a professional (eg, "like some kind of call in line. Like a hotline … where you could access a live expert…. something personal").

Likely Use of E-Mental Health Programs

Among the participants, 69.0% reported "yes" or "maybe" to potentially using an e-mental health program to deal with work stress. The percentage was higher in the high-risk group (72.6%), those in the levels of higher education and personal income, younger age groups, and those working in mid and large companies, compared with their counterparts. No differences were found by language, marital status, and employment status (employee vs self-employed). High-MDE-risk participants who reported that they would not use an e-mental health program for dealing with stress were asked "what would make it difficult to use an e-health program?". The reported barriers included perceived stigma associated with accessing e-mental health support (eg, "…social stigma, comfort of access," "….workplace ignorance and what do they call that where you stereotype ..."), lack of personal interaction inherent to e-mental health (eg, "lack of personal face to face," "… don't see the value of it if you could talk to your family doctor…"), lack of time (eg, "…if it was time consuming…"), and lack of knowledge (eg, "Well the fact that I don't know what an e health program is makes it difficult. I'm not sure that (laughs)").

Discussion

Principal Results

One key finding of this study was that 62.7% participants who were at high risk of having MDE had used the Internet for health information in the 12 months prior to the survey. This percent is slightly higher than a similar estimate from the 2012 Canadian Internet Use survey in which 60.8% men reported use of Internet for medical or health-related information [19]. Furthermore, more than 75% of high-MDE-risk men in our sample considered health information on the Internet to be useful in helping them

make health decisions and more than 72% would use an e-mental health program to deal with work related stress. Given that men often delay help-seeking for mental health problems because of perceived stigma and gender norms, our results suggest that the privacy inherent to e-mental health programs makes e-mental health programs a promising tool for improving men's mental health.

Acceptability of a tool is vital to evaluation of its effectiveness and implementation. Therefore, to develop e-mental health programs for men, it is critical to understand their preferred design features. It is enlightening to observe, from our survey, that "information about improving sleep hygiene" was the top design feature preferred by men, irrespective of their risk status. Individuals who are at high risk of MDE may be occupied by unhelpful thinking and look for strategies to solve the issues they encounter. Thus, it is not surprising that the second top feature they endorsed was "practice and exercise to help reduce symptoms of stress and depression" which is consistent with the principles of CBT, for example, changing unhelpful thinking and behaviors and problem-solving focus. We anticipated that CBT practices and educational information ("having access to quality information and resources about work stress issues") would be needed by the participants, and this was demonstrated in this study. This also is consistent with the fact that most of the existing e-mental health programs, such as MoodGYM [20], were developed based on the CBT approach [21]. We found that men who were at high risk of having an MDE were more likely to have endorsed the design features than men who were at low risk. No age differences were found in preferred design features. English-speaking participants were more likely to use CBT techniques and an app and French-speaking participants were more likely to use mood-monitoring tools. These results indicate that e-mental health programs incorporating these preferred features are likely to be used by men who are at high risk of an MDE across age and English- or French-speaking categories.

Comparison With Prior Work

Understanding the barriers to the use of e-mental health programs is also important for the development, evaluation, and implementation of the programs. Some features preferred by the participants reflect the concerns they have about e-mental health programs and potential barriers to the use. Based on our quantitative and qualitative data, confidentiality and privacy are the prominent concerns for high-risk participants. They were concerned about the consequences if others know that they use

the program to deal with stress related issues (eg, perceived stigma). Other barriers include extent to which the program is easy to use and navigate, credible (eg, information is provided by health professionals), relevant to one's personal situation, and interactive (eg, being able to communicate with a professional). Additionally, lack of time and knowledge about e-health are potential barriers reported by the participants. Schneider et al [22] investigated users' views of an online CBT program (MoodGYM) in a randomized controlled trial. Wetterlin et al's [23] cross-sectional study examined youth expectations for mental health websites. Both studies reported preferences and perceived barriers that are consistent with the results of our survey.

Limitations

Our study has several limitations. First, the survey data relied on self-report. Therefore, reporting and recall biases are possible. Second, our target population is Canadian working men who were aged 18 and older. As compared to men in the Canadian workforce in 2014, our sample was slightly older. The proportion of participants aged 18 to 29 in this study was 12.1%, whereas it was 20.2% in the Canadian workforce. Therefore, the proportions of some design features by age groups could have been overestimated or underestimated because of potential selection bias. Given the increasing use of mobile phones in young adults, future studies may investigate strategies for recruiting young adults through mobile phones. Finally, our survey collected self-reported qualitative information about barriers to the use of e-mental health programs. The qualitative information should be considered preliminary. More studies are needed to provide definitive answers.

Conclusions

There is a pressing need for developing innovative strategies for prevention of depression in men. This is a challenging endeavor, given the gender norms and social stigma against depression and help seeking among men. E-mental health holds potential as it can be confidential, easily accessible, and economic if designed appropriately. More studies are needed to examine preferred design features and the barriers to use in different populations so that e-mental health strategies that meet the needs of different age groups and personal backgrounds can be developed.

References

1. Patten SB, Williams JV, Lavorato DH, Wang JL, McDonald K, Bulloch Andrew G M. Descriptive epidemiology of major depressive disorder in Canada in 2012. Can J Psychiatry 2015 Jan;60(1):23-30 [FREE Full text] [Medline: 25886546]

2. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, et al. The Epidemiology of Major Depressive Disorder Results From the National Comorbidity Survey Replication (NCS-R). JAMA 2003;289(23):3095-3105. [doi: 10.1001/jama.289.23.3095]

3. Ferrari AJ, Charlson FJ, Norman RE, Patten SB, Freedman G, Murray CJL, et al. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. PLoS Med 2013 Nov;10(11):e1001547 [FREE Full text] [doi: 10.1371/journal.pmed.1001547] [Medline: 24223526]

4. Stewart WF, Ricci JA, Chee E, Hahn SR, Morganstein D. Cost of lost productive work time among US workers with depression. JAMA 2003 Jun 18;289(23):3135-3144. [doi: 10.1001/jama.289.23.3135] [Medline: 12813119]

5. Smith ND, Kawachi I. State-level social capital and suicide mortality in the 50 U.S. states. Soc Sci Med 2014 Nov;120:269-277. [doi: 10.1016/j.socscimed.2014.09.007] [Medline: 25262314]

6. Statistics Canada. 2014. Suicide rates, an overview, 1950 to 2009 URL: http://www.statcan.gc.ca/access_acces/archive.action?loc=/daily-quotidien/120725/dq120725a-eng.pdf&archive=1 [accessed 2016-05-05] [WebCite Cache ID 6hHsFZUt3]

7. Wang JL, Patten SB, Currie S, Sareen J, Schmitz N. A population-based longitudinal study on work environmental factors and the risk of major depressive disorder. Am J Epidemiol 2012 Jul 1;176(1):52-59 [FREE Full text] [doi: 10.1093/aje/kwr473] [Medline: 22556191]

8. Wang JL, Schmitz N, Dewa C, Stansfeld S. Changes in perceived job strain and the risk of major depression: results from a population-based longitudinal study. Am J Epidemiol 2009 May 1;169(9):1085-1091 [FREE Full text] [doi: 10.1093/aje/kwp037] [Medline: 19318611]

9. Wang JL, Lesage A, Schmitz N, Drapeau A. The relationship between work stress and mental disorders in men and women: findings from a population-based study. J Epidemiol Community Health 2008 Jan;62(1):42-47. [doi: 10.1136/jech.2006.050591] [Medline: 18079332]

10. Addis ME, Mahalik JR. Men, masculinity, and the contexts of help seeking. Am Psychol 2003 Jan;58(1):5-14. [Medline: 12674814]

XSL•FO

RenderX

11. NHS Confederation. E-Mental Health: What's all the Fuss About? 2014 URL: http://www.nhsconfed.org/~/media/Confederation/Files/Publications/Documents/E-mental-health.pdf [accessed 2016-03-22] [WebCite Cache ID 6gCakiVcz]

12. Cuijpers P, Donker T, van SA, Li J, Andersson G. Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. Psychol Med 2010 Dec;40(12):1943-1957. [doi: 10.1017/S0033291710000772] [Medline: 20406528]

13. Christensen H, Petrie K. State of the e-mental health field in Australia: where are we now? Aust N Z J Psychiatry 2013 Feb;47(2):117-120. [doi: 10.1177/0004867412471439] [Medline: 23297367]

14. Lal S, Adair CE. E-mental health: a rapid review of the literature. Psychiatr Serv 2014 Jan 1;65(1):24-32. [doi: 10.1176/appi.ps.201300009] [Medline: 24081188]

15. Andrews G. Implications for intervention and prevention from the New Zealand and Australian mental health surveys. Aust N Z J Psychiatry 2006 Oct;40(10):827-829. [doi: 10.1111/j.1440-1614.2006.01900.x] [Medline: 16959007]

16. Wang JL, Manuel D, Williams J, Schmitz N, Gilmour H, Patten S, et al. Development and validation of prediction algorithms for major depressive episode in the general population. J Affect Disord 2013 Oct;151(1):39-45 [FREE Full text] [doi: 10.1016/j.jad.2013.05.045] [Medline: 23790813]

17. D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation 2008 Feb 12;117(6):743-753 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.107.699579] [Medline: 18212285]

18. Lloyd-Jones DM. Cardiovascular risk prediction: basic concepts, current status, and future directions. Circulation 2010 Apr 20;121(15):1768-1777 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.109.849166] [Medline: 20404268]

19. Statistics Canada. 2013. Canadian Internet Use Survey URL: http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&Id=130941 [accessed 2016-03-26] [WebCite Cache ID 6gCZ4Z4aW]

20. Christensen H, Griffiths KM, Korten A. Web-based cognitive behavior therapy: analysis of site usage and changes in depression and anxiety scores. J Med Internet Res 2002;4(1):e3 [FREE Full text] [doi: 10.2196/jmir.4.1.e3] [Medline: 11956035]

21. Gratzer D, Khalid-Khan F. Internet-delivered cognitive behavioural therapy in the treatment of psychiatric illness. CMAJ 2016 Mar 1;188(4):263-272. [doi: 10.1503/cmaj.150007] [Medline: 26527829]

22. Schneider J, Sarrami FP, Grime P, Thornicroft G. Acceptability of online self-help to people with depression: users' views of MoodGYM versus informational websites. J Med Internet Res 2014;16(3):e90 [FREE Full text] [doi: 10.2196/jmir.2871] [Medline: 24681717]

23. Wetterlin FM, Mar MY, Neilson EK, Werker GR, Krausz M. eMental health experiences and expectations: a survey of youths' Web-based resource preferences in Canada. J Med Internet Res 2014;16(12):e293 [FREE Full text] [doi: 10.2196/jmir.3526] [Medline: 25519847]

XSL•FO
RenderX

Original Paper

Barriers and Facilitation Measures Related to People With Mental Disorders When Using the Web: A Systematic Review

Renaldo Bernard[1], MSc; Carla Sabariego[1], PhD; Alarcos Cieza[2], PhD

[1]Department of Medical Informatics, Biometry and Epidemiology – IBE, Chair for Public Health and Health Services Research, Research Unit for Biopsychosocial Health, Ludwig-Maximilians-Universität München, Munich, Germany

[2]Blindness and Deafness Prevention, Disability and Rehabilitation (BDD), World Health Organization, Geneva, Switzerland

Corresponding Author:
Renaldo Bernard, MSc
Department of Medical Informatics, Biometry and Epidemiology – IBE
Chair for Public Health and Health Services Research, Research Unit for Biopsychosocial Health
Ludwig-Maximilians-Universität München
Marchioninistraße 17
Munich, 81377
Germany
Phone: 49 89 2180 78229
Fax: 49 89 2180 78230
Email: renaldo.bernard@med.lmu.de

Abstract

Background: Mental disorders (MDs) affect almost 1 in 4 adults at some point during their lifetime, and coupled with substance use disorders are the fifth leading cause of disability adjusted life years worldwide. People with these disorders often use the Web as an informational resource, platform for convenient self-directed treatment, and a means for many other kinds of support. However, some features of the Web can potentially erect barriers for this group that limit their access to these benefits, and there is a lack of research looking into this eventuality. Therefore, it is important to identify gaps in knowledge about "what" barriers exist and "how" they could be addressed so that this knowledge can inform Web professionals who aim to ensure the Web is inclusive to this population.

Objective: The objective of this study was to provide an overview of existing evidence regarding the barriers people with mental disorders experience when using the Web and the facilitation measures used to address such barriers.

Methods: This study involved a systematic review of studies that have considered the difficulties people with mental disorders experience when using digital technologies. Digital technologies were included because knowledge about any barriers here would likely be also applicable to the Web. A synthesis was performed by categorizing data according to the 4 foundational principles of Web accessibility as proposed by the World Wide Web Consortium, which forms the necessary basis for anyone to gain adequate access to the Web. Facilitation measures recommended by studies were later summarized into a set of minimal recommendations.

Results: A total of 16 publications were included in this review, comprising 13 studies and 3 international guidelines. Findings suggest that people with mental disorders experience barriers that limit how they perceive, understand, and operate websites. Identified facilitation measures target these barriers in addition to ensuring that Web content can be reliably interpreted by a wide range of user applications.

Conclusions: People with mental disorders encounter barriers on the Web, and attempts have been made to remove or reduce these barriers. As forewarned by experts in the area, only a few studies investigating this issue were found. More rigorous research is needed to be exhaustive and to have a larger impact on improving the Web for people with mental disorders.

XSL•FO
RenderX

Introduction

Mental disorders (MDs) are a significant public health issue owing to their high impact on people with these disorders, in terms of restrictions placed on their participation in all areas of life, family life and the wider society. Mental disorders affect almost 1 in 4 adults at some point during their lifetime [1] and coupled with substance use disorders are the fifth leading cause of disability adjusted life years worldwide [2]. People with mental disorders (PwMD) often experience similar impairments, activity limitations, and restricted participation in life events, even with the diversity in symptoms and etiology associated with these conditions [3]. Family members often provide care, which sometimes puts a strain on familial relationships, reduces opportunities for leisure, and negatively impacts finances due to time spent providing care instead of working [4]. The associated reduction in productivity from both affected persons and their family can translate to a decrease in contributions to the local economy [5]. In addition, having a large segment of the population subscribing to treatment and support services incurs considerable costs [5].

The Web is often used as a source of support for PwMD and shows great promise for the reduction of the burden of mental disorders. Mental health–related Web browsing, primarily for information seeking, is common among PwMD [6,7]. Web-based mental health communities are known to supplement traditional mental health services [8] and act as an important factor in encouraging PwMD to seek professional help [9]. A recent meta-analysis has indicated that guided Web-based cognitive behavioral therapy may be as effective as the face-to-face equivalent for social anxiety disorder, panic disorder, spider phobia, and depressive symptoms [10]. Many other Web-based treatment and intervention options are increasingly being explored for other mental disorders (eg, posttraumatic stress disorder, eating disorders) [11] and populations including children (eg, Project CATCH-IT, MoodGYM) [12,13] with positive results.

There are also features of the Web environment that could potentially limit how much PwMD who experience cognitive deficits can benefit from the Web. Using the Web is considered a very cognitively demanding activity requiring not only good knowledge and understanding of Web features (eg, search engines) but also the ability to quickly analyze, synthesize, evaluate, and apply presented information while avoiding inconsequential details (eg, adverts and untrustworthy information) that are abundant on the Web [14]. Several cognitive domains, including executive functioning, attention, and memory, are commonly impaired in PwMD [15]. These impairments may be linked to difficulties using the Web such as when performing Web searches, task switching, retaining and recalling information, and ignoring distractions (eg, adverts) to focus attention. Moreover, the Web has also been found to be relatively absent of nonverbal and social context cues (eg, gestures, facial expression) compared with off-line [16,17]. These cues are important for guiding behavior when interacting with others, and their absence could make social interaction difficult. Although Web users are normally able to skillfully compensate and overcome these "deficiencies" [18], sometimes

even by capitalizing on them [19], it could be challenging for PwMD who experience cognitive deficits to do the same.

People with mental disorders have received little attention from Web accessibility research despite increased inquiries into the difficulty others with cognitive impairment face on the Web. This research gap was highlighted over a decade ago [20,21], and more recently, there has been some indication that the gap still exists [22]. Current recommendations also prescribe the same treatment to address accessibility for PwMD and a myriad of other diverse conditions that fall under the broad heading of conditions associated with cognitive limitations (eg, intellectual disabilities, multiple sclerosis) [23].

A comprehensive review of literature concerned with the barriers PwMD encounter when using the Web and/or the facilitation measures developed to address these barriers is needed to ensure that the Web is inclusive to this population. Available knowledge will support Web professionals in making well-informed choices about the removal of barriers affecting PwMD. If this is not possible, it may instead provide facilitation measures to accommodate this group. As a result, Web-based resources could be systematically evaluated for compliance with measures that are known to remove barriers or provide facilitation for PwMD. Identified gaps in knowledge about "what" barriers exist and "how" they could be addressed—based on a comparison and integration of what is known on the topic—is likely to encourage further research into these highlighted areas as well.

The objective of this systematic review was to provide an overview of the existing evidence regarding the barriers PwMD experience when using the Web and facilitation measures used to address such barriers. Specific aims are to detail barriers and facilitation measures, how they were identified or developed, and related trends (ie, the extent of coverage for specific mental disorders or digital technologies, study designs used, publication recency, and research region).

Methods

A systematic review was carried out to identify barriers PwMD encounter when using the Web and the recommended facilitation measures to remove or reduce these barriers.

Search Strategy

Search terms were broadly based on concepts relating to Web accessibility, mental disorders, and also digital technologies (see Multimedia Appendix 1). Digital technologies were included because knowledge about any barriers here would likely be also applicable to the Web. This was also a proactive measure to avoid having the review suffer from the paucity of research in the area as revealed by preliminary searches. Databases searched include MEDLINE, PsycARTICLES, CINAHL, Library, Information Science &Technology Abstracts, Computers & Applied Sciences Complete, Inspec, Web of Science Core Collection. Reference lists of included publications were also searched to avoid missing relevant publications not identified during the search of databases. There were no publication date restrictions to ensure that the review included as many studies as possible. There was also no restriction to

empirical studies. Other types of publications such as international standards and guidelines are usually widely adopted and highly regarded and can be especially helpful when there is insufficient empirical evidence on a particular issue.

Eligibility Criteria

Included publications describe the difficulties PwMD encounter when using any digital technology or provide guidance on how to improve the accessibility of any digital technology for this group. All mental disorders were considered regardless of a formal diagnosis or not. All digital technologies such as computers, video games, mobile devices, and websites were also considered. Journal articles, gray literature, international and national standards and guidelines, reports, and conference proceedings written in the English language were considered for inclusion. Publications in the form of commentaries, letters to the editors, and editorials were excluded.

Eligibility Assessment

One reviewer (RB) screened all abstracts, and another (DH) screened 84% (1692/2013) selected at random. Both screenings were conducted independently to reduce the chance of reviewer bias and increase reliability [24]. Inconsistences in ratings—eligible, ambiguous, or excluded—were later discussed and resolved by consensus. One reviewer (RB) then appraised the full texts of abstracts rated as eligible.

Data Extraction and Synthesis of Results

Information extracted from studies was study characteristics—publication year, country, study design, methods and participants or target population (eg, mental disorders, age, gender, and education); barriers and facilitation measures—process used for the development of the facilitation measure and related mental disorders; and definitions of accessibility and disability. Data extracted from other documents—international standards and guidelines—did not include information about study designs and participants (eg, age and gender).

The International Classification of Functioning, Disability and Health was used to define barriers and facilitation measures [25]. Factors (eg, small font, complicated language) that through their absence or presence limit functioning were identified as barriers. Conversely, factors (eg, legible font, simple language) that instead improve functioning through their absence or presence were identified as facilitation measures.

Synthesis was performed by categorizing all findings and later summarizing facilitation measures recommended by studies. Data were first categorized according to the 4 foundational principles of Web accessibility: operable—user interface components and navigation must be easy and safe to use; understandable—information and the operation of a user interface must be easily interpreted accurately; perceivable—information and user interface components must be presentable to users in ways they can be sufficiently aware on these components; robust—content must be flexible enough that a wide range of user agents, including technologies that enable persons with disabilities to perform tasks that would be otherwise challenging (ie, assistive technologies), can interpret it reliably [26]. These 4 foundational principles were proposed by the World Wide Web Consortium (W3C) and form the necessary basis for anyone to gain adequate access to the Web. Results from studies came from 2 sources—expert opinion or empirical research—and they are labeled to denote these different sources. Facilitation measures from guidelines are also labeled for easy identification. Facilitation measures recommended by studies were later summarized into a set of minimal recommendations after the categorization of findings. Those from guidelines have already been aptly summarized elsewhere [27-29].

Results

A total of 16 publications were included in this review, comprising 13 studies reporting on the usability of various technologies [30-40] and Internet or computer use among PwMD [41,42] and 3 international guidelines [23,43,44], which were all developed by the W3C. These guidelines have been adopted by many governments and are also widely considered as the international standard for Web accessibility. A flow chart of the review process is presented in Figure 1.

Figure 1. Flowchart of the review identification and selection process.
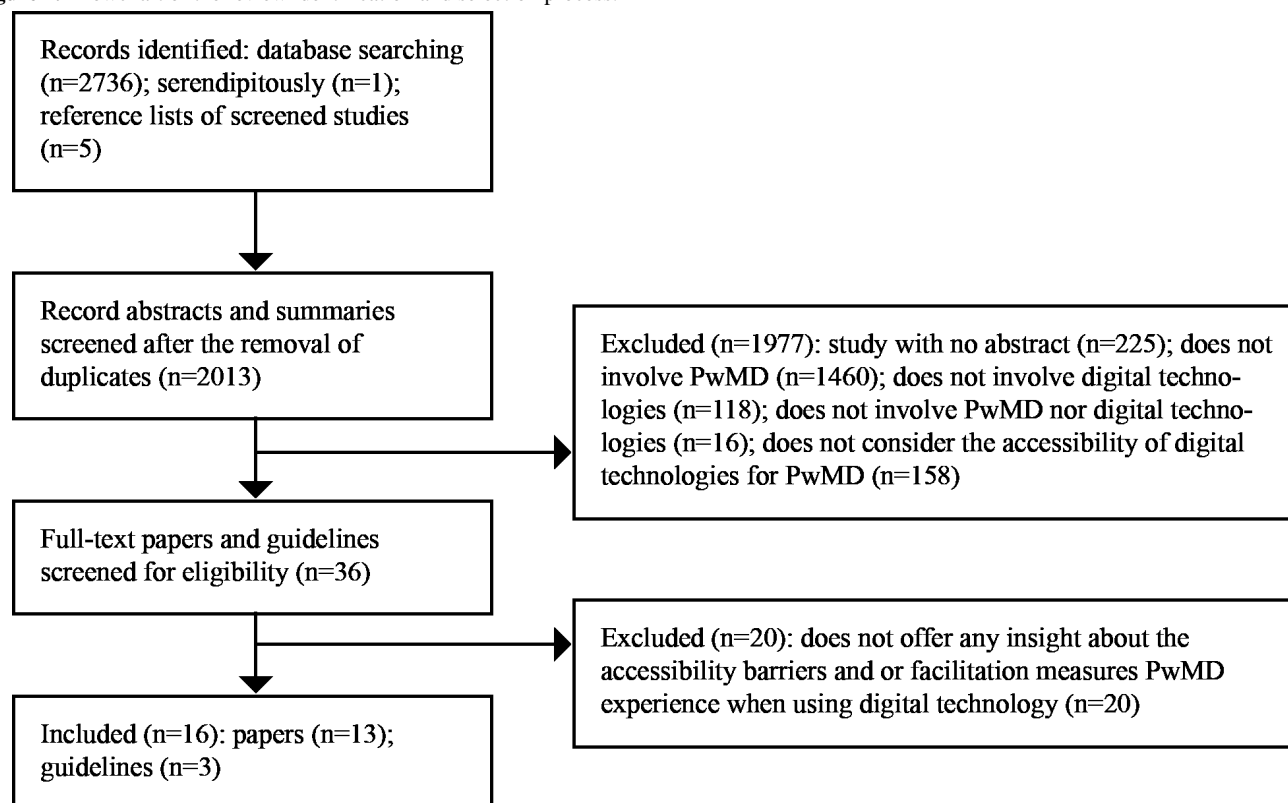
Study and Guideline Characteristics

Nine of the included studies [30-33,35,37-39,42] originated in the United States, 2 studies [34,40] in the United Kingdom, one [41] in Austria, and another [36] in Sweden as summarized in Table 1. Over 62% (10/16) of the included publications [32-34,37,38,42] were published within the last 5 years, and the earliest [35] was published in 1998.

All 3 included guidelines were published by the W3C based in the United States. However, the guidelines are the result of collaboration among international experts. Two of the three included guidelines (User Agent Accessibility Guidelines 1.0 and Authoring Tool Accessibility Guidelines 1.0) were published over 12 years ago, and the third (Web Content Accessibility Guidelines 2.0) was published in 2008.

Table 1. Characteristics of included publications.

Citation, sample size (n), and year	Origin country	Study design	Digital technology	Diagnosis
[35], 52, 1998	United States	Qualitative, focus groups and interviews, clustering and summation	Multimedia application	Depression
[43], 2000	United States	Guideline	Web	Mental disorders
[31], 5, 2002	United States	Qualitative, focus group and usability testing, content analysis	Website	Anxiety disorder and depression
[44], 2002	United States	Guideline	Web	Mental disorders
[38], 98, 2007	United States	Quantitative, usability testing	Website	Bipolar disorder, schizophrenia, schizoaffective disorder, depression
[23], 2008	United States	Guideline	Web	Mental disorders
[41], 26, 2010	Austria	Qualitative, interviews, content analysis	Internet and website	Schizophrenia, schizoaffective disorder[a]
[32] (n=16), 2011	United States	Mixed, interviews, usability testing and expert review, thematic analysis, and descriptive statistics	Website	Schizophrenia, bipolar disorder, depression
[33], 71, 2011	United States	Qualitative, interviews and usability Testing, descriptive statistics, and *t*-tests	Website	Severe mental illness
[39], 149, 2012	United States	Quantitative (fractional factorial experimental design), usability testing, polychotomous logistic regression, and mixed-effect regression	Website	Substance use disorder, schizophrenia, depression, bipolar disorder, other psychotic disorder, schizoaffective disorder, anxiety disorder[b]
[42], 28, 2013	United States	Qualitative, interviews and observations, thematic and task analysis	Computer and website	Schizophrenia, bipolar disorder, depression, anxiety disorder, schizoaffective disorder
[37], 38, 2013	United States	Quantitative, usability testing, linear mixed-effect regression	Website	Schizophrenia, schizoaffective disorder[b]
[40], 12, 2013	United Kingdom	Qualitative, focus group, thematic analysis	Website	Bipolar disorder[b]
[30], 924, 2013	United States Mixed, usability testing and survey, thematic analysis, descriptive statistics	Mobile phone and website	Schizophrenia, schizoaffective disorder	
[34], 20, 2014	United Kingdom	Qualitative, focus group, thematic analysis	Website	Depression, anxiety disorder
[36], ≥100, 2015	Sweden	Qualitative, focus group, thematic analysis Digital technologies	Bipolar disorder, depression, schizophrenia, anxiety disorder, mental disorders	

[a] Diagnosis was established using the International Classification of Diseases, 10th revision.

[b] Diagnosis was established using the Diagnostic and Statistical Manual of Mental Disorders, 4th edition.

Design and Methods

Nine of the included studies investigated the usability of Web-based resources [30-34,37-40] and multimedia tools [35]. One study focused on Internet use [41], one on the use of digital technologies [36], one on the development of a mobile phone system [30], and another on computer use [42] among PwMD. Eight of the included studies used qualitative methods [31,33-36,40-42], 3 [37-39] adopted a quantitative approach, and 2 [30,32] used mixed methods. Seven studies used usability testing [30-33,37-39], 5 used interviews [32,33,35,41,42], 5 used focus groups [32,33,35,41,42], and single studies used observations [42], survey [30], and user testing.

The 3 included guidelines [23,43,44] were primarily developed based on contributions over several years from experts involved in international working groups on varying aspects of Web accessibility [45].

XSL•FO
RenderX

Sample Characteristics

Sample sizes for included studies ranged from 5 to >100 (mean 48). Overall, 11 studies [30-33,36-42] reported the age of participants, which ranged from 18 to at least 75 years. Three studies [37,39,40] used the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM IV), 1 [41] used the International Classification of Diseases, 10th revision (ICD-10), and the remaining studies did not mention the use of a classification of mental disorders. Samples including people with schizophrenia (69%) [30,32,35-39,41,42] were most common among the 13 included studies, followed by samples where participants were affected by depression (62%) [31,32,34-36,38,39,42], schizoaffective disorder (46%) [30,37-39,41,42], anxiety disorders (38%) [31,34,36,39,42], and bipolar disorder (38%) [32,36,38,40,42]. Single studies reported that participants had severe mental illness (SMI) (eg, schizophrenia, schizoaffective disorder, bipolar disorder, and major depression) [33], mental disorders [36], psychotic disorders [39], and substance use disorder [39] but did not state any particular mental disorder. Most studies considered more than 1 mental disorder except [33], which focused on schizophrenia and [40] on bipolar disorder.

All 3 included guidelines were developed to give guidance on how to remove and reduce barriers experienced by people with a range of disabilities including auditory, cognitive, and neurological, physical, speech, and visual disabilities. Extracted guidelines were identified by the authors of the guidelines as being relevant to cognitive and neurological disorders [46]. These disorders include attention-deficit hyperactivity disorder, autism spectrum disorder, intellectual disabilities, learning disabilities, memory impairments, multiple sclerosis, perceptual disabilities, seizure disorders, and mental disorders. No particular mental disorder was specified.

Digital Technology

As summarized in Table 1, websites were the most studied digital technology, followed by single studies each investigating either computers [42] or multimedia applications [35]. Only three studies [42,30,36] investigated more than 1 technology, viz computers and websites, mobile phone and websites, and several digital technologies, respectively. The 3 included guidelines target websites (ie, Web Content Accessibility Guidelines 2.0), user agents (ie, any software that retrieves, renders, and facilitates end user interaction with Web content; User Agent Accessibility Guidelines 1.0) and Web authoring tools (Authoring Tool Accessibility Guidelines 1.0).

Scope of Barriers and Facilitation Measures Related to Digital Technology Usage by PwMD

Included studies revealed 42 barriers and 59 facilitation measures. These are summarized in Tables 2 and 5. Four studies [31,32,35,37] did not mention any barriers and 2 [36,41] no facilitation measures. Four studies [30,33,34,38] recommended facilitation measures to address barriers, and only 25 of these pairings were identified.

The 3 included guidelines recommended 30 facilitation measures and did not explicitly report any barriers. However, the W3C has published several barriers on its website that people with cognitive and neurological disabilities including mental health disabilities face when using the Web. Examples of these barriers include complex navigation mechanisms, page layouts that are difficult to understand and use, and moving, blinking, or flickering content, and background audio that cannot be turned off [46].

Of the 131 identified barriers and facilitation measures, 63 were relevant to depression (48%), 54 to schizophrenia (41%), 48 to anxiety disorders (37%), 39 to bipolar disorder (30%), 37 to mental disorders (28%), 35 to schizoaffective disorder (27%), 11 to SMI (8%), and 3 to substance abuse and psychotic disorders equally (2%). Most of the 42 identified barriers were relevant to people with depression (64%), followed by those with an anxiety disorder (62%), schizophrenia (50%), bipolar disorder (40%), schizoaffective disorder (31%), mental disorders (17%), SMI (12%), and substance use disorder and other psychotic disorders equally (2%). Identified facilitation measures (n=89) mostly targeted people with depression (40%), schizophrenia (37%), mental disorders (34%), and anxiety, bipolar disorder and schizoaffective disorder equally (25%). SMI (7%) and substance use disorder and other psychotic disorders equally (2%) accounted for a small portion of the identified facilitation measures.

All barriers identified were revealed by research findings. Identified facilitation measures were proposed directly from research findings (n=31) [30,33,37-40], by international working groups of experts in the area of accessibility (n=30) [23,43,44] and expert opinion of researchers conducting studies (n=28) [31,32,34,35,42].

Table 2. Barriers and facilitation measures categorized by the 'perceivable' foundational principle of Web accessibility.

Barrier	Facilitation measure
Unable to locate information [34]	Provide intuitive navigation and ensure information filters and search functions work properly[a].
Nonperceivable icons [34]	Avoid complicated language and ensure menu options and links are easy to understand[a].
Too small font [30]	Increase font size[b].
Difficulty reading small font and with eye strain [42]	Use small but legible font and refrain from using graphics in websites with shallow information hierarchies that do not feature navigational lists[b] [39].
	Use large navigation buttons[a] [32].
	Use a minimal number of colors that differentiates information and contrasts well[a] [31].
	Use a simple design with pages that are pleasing to the eye and easy to read[a] [31].
	Use graphics that are purposeful to the website[a] [31].
	Prominently present hyperlinks: ensure clear labeling and highly visible positioning[b] [37].
	Make hyperlinks' text as explicit as possible[b] [37].
	List hyperlinks for a given topic together in a single column[b] [37].
	Font size, buttons, and links should be sufficiently large to ensure usability[a] [42].
	Use attention grabbing and not boring design[b] [40].
	Guideline 1.1: Provide text alternatives for any nontext content so that it can be changed into other forms people need, such as large print, braille, speech, symbols, or simpler language[c] [23].
	Guideline 1.2: Provide alternatives for time-based media[c] [23].
	Guideline 1.3: Create content that can be presented in different ways (eg, simpler layout) without losing information or structure[c] [23].
	Guideline 1.4: Make it easier for users to see and hear content including separating foreground from background[c] [23].
	Guideline 5: Ensure that the user can control the behavior of viewports (ie, screen) and user interface controls, including those that may be manipulated by the author (eg, through scripts—list of computer commands)[c] [44].
	Guideline 3: Support the creation of accessible content[c] [43].
	Guideline 2: Generate standard markup (ie, document annotations)[c] [43].
	Guideline 1: Support accessible authoring practices[c] [43].
	Guideline 7: Ensure that the authoring tool is accessible to authors with disabilities[c] [43].
	Guideline 2: Ensure that users have access to all content, notably conditional content that may have been provided to meet the requirements of the Web Content Accessibility Guidelines 1.0[c] [44].
	Guideline 3: Ensure that the user may turn off rendering of content (eg, audio, video, scripts) that may reduce accessibility by obscuring other content or disorienting the user[c] [44].
	Guideline 4: Ensure that the user can select preferred styles (eg, colors, size of rendered text, and synthesized speech characteristics) from choices offered by the user agent. Allow the user to override author-specified and user agent default styles[c] [44].
	Guideline 11: Allow users to configure the user agent so that frequently performed tasks are made convenient and allow users to save their preferences[c] [44].

[a]Facilitation measure derived from expert opinion of researcher(s) conducting a study.

[b]Facilitation measure medived from empirical evidence.

[c]Facilitation measure derived from working group of experts.

Table 3. Barriers and facilitation measures categorized by the 'understandable' foundational principle of Web accessibility.

Barrier	Facilitation measure
Information overload [34]	Ensure information is organized well and avoids distracting design[a].
Poor organization and presentation [34]	Ensure information is organized well and avoids distracting design[a].
Excessive advertisements [34]	Ensure information is organized well and avoids distracting design[a].
Confusing menu options [34]	Avoid complicated language and ensure menu options and links are easy to understand[a].
Complicated language [34]	Avoid complicated language and ensure menu options and links are easy to understand[a].
Complex purchasing process [34]	Avoid complicated language and ensure menu options and links are easy to understand[a].
Distracting design [34]	Ensure information is organized well and avoids distracting design[a].
Use of abstract reasoning [38]	Present text at a low reading level[b].
Difficulty comprehending text [33]	Present text in large font and language below a fifth-grade reading level[b].
Difficulty understanding abbreviations [30]	Remove abbreviations[b].
Difficulty understanding long words [30]	Reduce text[b].
Too lengthy text [30]	Simplify wording to fourth-grade level[b].
Overabundance of information [41]	
Unwanted movements or flickering [36]	
Cluttered design [36]	
Lack of logic and consequence in concept and design [36]	Provide resources in video and audio format[a] [35].
	Use a modular and hierarchical approach when presenting information[a] [35].
	Present important information first[a] [35].
	Use large navigation buttons[a] [32].
	Provide explicit labels that use longer concrete phrases to describe content[a] [32].
	Explicit instructions on how to use the website[a] [32].
	Provide text at fifth-grade reading level[a] [32].
	Provide instructions on how to navigate programs and websites[a] [42].
	Use a simple design with pages that are pleasing to the eye and easy to read[a] [31].
	Provide category headings that clearly identify what information is underneath[a] [31]. Use menus with options that are ordered in a meaningful way and/or have an evident hierarchy[a] [31].
	Give a clear identity to the homepage[a] [31]. Provide a homepage with just the right amount of information (graphics, text, links) to make the page understandable without overwhelming the user[a] [31].
	Use language that the user can identify with[a] [31].
	Meaningfully group of information[a] [31].
	Use graphics that are purposeful to the website[a] [31].
	Comprehensively list hyperlinks surrounding a given topic[b] [37].
	Include minimal amount of content on pages[b] [37].
	Single topic of interest: group hyperlinks and topics in one area of the screen[b] [37].
	List hyperlinks for a given topic together in a single column[b] [37].
	Use an ample number of images and visual aids[b] [30].
	Provide content users can identify with (eg, case stories, worked examples, and success stories)[b] [40].

Barrier	Facilitation measure
	Use a flat hierarchy[b] [38].
	Provide explicit labeling[b] [38].
	Use lower-level modules (eg, code and data to implement a specific functionality)[b] [38].
	Use familiar phrasing[b] [38].
	Guideline 3.1: Make text content readable and understandable[c] [23].
	Guideline 3.2: Make Web pages appear and operate in predictable ways[c] [23].
	Guideline 3.3: Help users avoid and correct mistakes[c] [23].
	Guideline 7: Observe operating environment conventions for the user agent user interface, documentation, input configurations, and installation[c] [44].
	Guideline 12: Ensure that the user can learn about software features that benefit accessibility from the documentation. Ensure that the documentation is accessible[c] [44].
	Guideline 2: Ensure that users have access to all content, notably conditional content that may have been provided to meet the requirements of the Web Content Accessibility Guidelines 1.0[c] [44].
	Guideline 3: Ensure that the user may turn off rendering of content (eg, audio, video, scripts) that may reduce accessibility by obscuring other content or disorienting the user[c] [44].
	Guideline 4: Ensure that the user can select preferred styles (eg, colors, the size of rendered text, and synthesized speech characteristics) from choices offered by the user agent. Allow the user to override author-specified and user agent default styles[c] [44].
	Guideline 5: Integrate accessibility solutions into the overall "look and feel"[c] [43].
	Guideline 6: Promote accessibility in help and documentation[c] [43].
	Guideline 4: Provide ways of checking and correcting inaccessible content[c] [43].
	Guideline 1: Support accessible authoring practices[c] [43].
	Guideline 7: Ensure that the authoring tool is accessible to authors with disabilities[c] [43].
	Guideline 3: Support the creation of accessible content[c] [43].
	Guideline 2: Generate standard markup[c] [43].

[a]Facilitation measure derived from expert opinion of researcher(s) conducting a study.

[b]Facilitation measure derived from empirical evidence.

Table 4. Barriers and facilitation measures categorized by the 'operable' foundational principle of Web accessibility.

Barrier	Facilitation measure
Poor navigation [34]	Provide intuitive navigation[a].
Poor information filters [34]	Ensure filters and search functions work properly[a].
Information overload [34]	Ensure information is organized well and avoid distracting design[a].
Difficulty with fine motor coordination [33]	Change double clicking to single clicking[c].
Difficulty clicking small radio buttons [33]	Change small buttons to large buttons[b].
Difficulty using a mouse [33]	Create video mouse tutorial[b].
Lack of knowledge on how to navigate a website [33]	Create basic instructions on how to change screens[b].
Lack of knowledge on how to navigate a website [33]	Create a flat website (without multiple layers)[b].
Too close and sensitive touchscreen buttons [30]	Enlarge buttons and space between them and require long enough touch-and-release functionality[b].
Navigating a website with more than 5 hierarchical levels [39]	Use 99 words or less, 2 navigational areas or less, 7 hyperlinks or less, and few topic areas covered per page and no graphics and toolbars[b].
Time-limited response forms [34]	
Slow response in websites loading information [34]	
Necessity to distance oneself from illness-related topics as part of the recovery process [41]	
Difficulty operating a computer mouse [42]	
Difficulty typing words in designated areas [42]	
Difficulty scrolling or using menu options to access information [42]	
Difficulty navigating [42]	
Processing delays [40]	
Broken links [40]	
Additional software requirements [40]	
Unwanted movements or flickering [36]	
Cluttered design [36]	
Evil design (when design is used to persuade or trick you to do something) [36]	
Functions and services with login [36]	
Lack of logic and consequence in concept and design [36]	
Lack of trustworthiness [36]	
Managing passwords and other codes (eg, Completely Automated Public Turing test to tell Computers and Humans Apart—CAPTCHA) [36]	Use a website with no more than 3 hierarchal levels and words per hyperlink and that has navigational lists[b] [39]. Use small but legible font and refrain from using graphics in websites with shallow hierarchies that do not feature navigational lists[b] [39]. Use of different media and technological additions (eg, reward logo or bookmark functionality)[b] [40]. Ensure resource can be easily used by people with low computer literacy[b] [40]. Allow users to progress through the system at their own pace[a] [35]. Pop-up menus that appear with hovering to reduce need for clicking[a] [32]. Use a shallow hierarchy (reach the destination within 2 clicks)[a] [32].

Barrier	Facilitation measure
	Use large navigation buttons[a] [32].
	Provide several options (eg, mouse, keyboard arrows, touch screen) to assist users when navigating programs and websites[a] [42].
	Provide instructions on how to navigate programs and websites[a] [42].
	Use shorter pages that do not require a lot of scrolling, especially for the home page[a] [31].
	Allow for personalization or getting the best fit[b] [40].
	Guideline 2.2: Provide users enough time to read and use the content[c] [23].
	Guideline 2.3: Do not design content in a way that is known to cause seizures[c] [23].
	Guideline 2.4: Provide ways to help users navigate, find content, and determine where they are[c] [23].
	Guideline 9: Provide access to content through a variety of navigation mechanisms, including sequential navigation, direct navigation, searches, and structured navigation[c] [44].
	Guideline 10: Provide information that will help the user understand browsing context[c] [44].
	Guideline 1: Ensure that the user can interact with the user agent (and the content it renders) through different input and output devices[c] [44].
	Guideline 5: Ensure that the user can control the behavior of viewports and user interface controls, including those that may be manipulated by the author (eg, through scripts)[b] [44].
	Guideline 2: Ensure that users have access to all content, notably conditional content that may have been provided to meet the requirements of the Web Content Accessibility Guidelines 1.0[b] [44].
	Guideline 3: Ensure that the user may turn off rendering of content (eg, audio, video, scripts) that may reduce accessibility by obscuring other content or disorienting the user[c] [44].
	Guideline 4: Ensure that the user can select preferred styles (eg, colors, the size of rendered text, and synthesized speech characteristics) from choices offered by the user agent. Allow the user to override author-specified and user agent default styles[c] [44].
	Guideline 7: Ensure that the authoring tool is accessible to authors with disabilities[c] [43].
	Guideline 1: Support accessible authoring practices[c] [43].
	Guideline 3: Support the creation of accessible content[c] [43].

[a]Facilitation measure derived from expert opinion of researcher(s) conducting a study.

[b]Facilitation measure derived from empirical evidence.

[c]Facilitation measure derived from working group of experts.

Table 5. Barriers and facilitation measures categorized by the 'robust' foundational principle of Web accessibility.

Barrier	Facilitation measure
	Guideline 3: Support the creation of accessible content[a] [43].
	Guideline 2: Generate standard markup[a] [43].
	Guideline 4: Provide ways of checking and correcting inaccessible content[a] [43].
	Guideline 6: Implement interoperable interfaces to communicate with other software (eg, assistive technologies, the operating environment, and plug-ins)[a] [44].
	Guideline 8: Support the accessibility features of all implemented specifications. Implement W3C Recommendations when available and appropriate for a task[a] [44].
	Guideline 7: Observe operating environment conventions for the user agent user interface, documentation, input configurations, and installation[a] [44].
	Guideline 1: Ensure that the user can interact with the user agent (and the content it renders) through different input and output devices[a] [44].
	Guideline 4.1: Maximize compatibility with current and future user agents, including assistive technologies[a] [23].

[a]Facilitation measure derived from working group of experts.

Synthesis of Results

Categorization of Results by Foundational Principles of Web Accessibility

The identified barriers and facilitation measures were categorized according to the foundational principles of Web accessibility that was proposed by the W3C and are summarized in Tables 2 and 5 —additional tables organized by categories can be requested. Each identified barrier and facilitation measure was sorted into multiple categories if applicable. The barriers resulted in 3 categories as none were assigned to the robust category: operable (n=26); understandable (n=16); perceivable (n=4). The facilitation measures resulted into 4 categories: operable (n=35); understandable (n=49); perceivable (n=26); and robust (n=8).

Some studies paired a barrier with a corresponding facilitation measure, and other studies did not. The former was categorized based on the barrier, and the latter was categorized based on the specific barrier or facilitation measure that was not paired. Linking barriers that were not paired with a corresponding facilitation measure was beyond the scope of this review. A synthesis of Tables 2 and 5 is presented in the following section.

Operable

Identified barriers and facilitation measures (n=61) in this category gave most coverage to depression (49%), followed by bipolar disorder (43%), anxiety (41%), schizophrenia (39%), mental disorders (34%), schizoaffective disorder (20%), SMI (16%), and substance use disorder and other psychotic disorders equally (7%).

Barriers reported by included studies are primarily related to poorly designed navigational elements (eg, content filters), difficulties with fine motor coordination (eg, clicking small radio buttons, operating computer mouse, scrolling), poorly designed pages with time-limited response forms, too much information, and unoptimized components that contribute to slow webpage loading times.

Facilitation measures derived from empirical evidence gave guidance on design involving a reduction in the number of clicks needed to select options, an increase in buttons sizes, and websites that feature a shallow hierarchical structure and allows for personalization. Facilitation measures based on the expert opinion of researchers conducting studies suggest that websites should incorporate efficient content filters with intuitive navigation and permit users to browse at their pace.

Most facilitation measures recommended by the 3 included guidelines were focused on increasing users' control. This involved providing users with enough time, alternative methods and information presentation styles, and instruction to interact with content. Other measures recommended that authoring tools must be accessible, promote accessible practices, and support the creation of accessible content.

Understandable

Most of the 64 identified barriers and facilitation measures in this category addressed depression (61%), schizophrenia (45%), anxiety (41%), mental disorders (34%), schizoaffective disorder (31%), and bipolar disorder (27%). However, SMI (3%) received considerably less coverage, and no barriers and facilitation measures were recorded for substance use disorder and other psychotic disorders in this category.

Included studies revealed barriers that included the use of complicated and excessive content, distracting and confusing design, and complex and overindulgent website functions (eg, excessive advertising and complicated purchasing processes). Facilitation measures derived from empirical evidence heavily focus on increasing the clarity of website content by ensuring only necessary information is shared and provided at a low reading level with no abbreviations and unfamiliar phrasing. Facilitation measures based on expert opinion focus more on the presentation and organization of website content. For example, they recommend the usage of alternative information formats, explicit labels that use concrete sentences to describe content and instructions, organizing content by importance, and forming meaningful content groups.

Facilitation measures from the 3 included guidelines recommend ways to help make content readable and understandable by ensuring abbreviations are expanded, reading level is appropriate, and providing explanations for any jargon used among other things. It was also recommended that several features should be incorporated into Web authoring tools: accessibility solutions in the design, mechanisms to correct inaccessible content and those that support accessible authoring practices.

Perceivable

Most of the 30 identified barriers and facilitation measures in this category targeted people with mental disorders (40%), depression (33%), anxiety and schizophrenia equally (30%), schizoaffective disorder (27%), bipolar disorder (17%) substance use disorder, and other psychotic disorders (3%). No barriers and facilitation measures were recorded for SMI in this category.

Identified barriers point to difficulties with reading small font, recognizing icons, and locating information. Facilitation measures derived both from empirical evidence and the expert opinion of researchers conducting studies recommend that links and other navigational elements should be easily recognizable, and use of images must be purposeful.

Facilitation measures recommended by the 3 included guidelines were predominantly focused on providing alternative content options and personal configurations for content. Other measures, all originating from the Authoring Tool Accessibility Guidelines1.0, generally recommend that authoring tools and practices must be accessible and support the creation of accessible content.

Robust

This category only contains facilitation measures from 1 of the 3 included guidelines, and no barriers were identified. All identified facilitation measures target PwMD. Recommended facilitation measures largely promote compatibility between user agents, authoring tools and Web content, and assistive technologies. The suggested methods to do this involve providing ways of checking and correcting inaccessible content within authoring tools and mainly adhering to standard markup, relevant W3C recommendations, and operating environment conventions.

Summary of Facilitation Measures Recommended by Studies

Facilitation measures recommended by studies were summarized into a group of 20 from 59 recommendations and are summarized in Table 6. Table 6 does not list or arrange summarized facilitation measures in any particular order. Nine of the summarized facilitation measures were the result of empirical work and 11 from the expert opinion of researchers.

Table 6. Summary of facilitation measures recommended by studies.

Derived from empirical evidence	Derived from expert opinion of researcher(s)
Provide instructions on how to change between different page views.	Provide intuitive navigation and ensure information filters and search functions work.
Build websites with a minimal number of layers.	Provide explicit instructions on how to use the website.
Provide legible font and perceivable buttons and links.	Use simple and familiar language with no abbreviations.
Comprehensively list hyperlinks surrounding a given topic.	Allow users to progress through the system at their own pace.
Allow for personalization or getting the best fit for the user.	Use graphics and colors sparingly and meaningfully.
Use of different media and technological additions (eg, reward logo or bookmark functionality).	Provide several options (eg, mouse, keyboard arrows, touch screen) to assist users with navigation.
Use attention-grabbing and not boring design.	Provide resources in video and audio format.
Use simple and familiar language.	Use legible font and sufficiently large buttons
Use an ample number of images and visual aids.	Use a simple design with webpages that are pleasing to the eye and easy to read.
	Meaningfully group information.
	Use a minimal amount of content.

Discussion

Principal Findings and Comparison With Prior Work

The 13 studies that could be included in this review support preexisting views [20,21] that there is little research on the barriers PwMD experience when using digital technology and facilitation measures used to address such barriers. Despite being few, included studies and guidelines give valuable insight into what is known and where knowledge gaps lie.

Barriers People With Mental Disorders Encounter When Using Digital Technologies

People with mental disorders encounter a wide range of barriers when using the Web that makes it difficult for them to perceive, understand, and operate this tool along with content contained therein. Most barriers result from distracting and confusing design, complicated content and website functions, an overabundance of information, and a high-demand for good fine-motor skills and rapid information processing. Persons affected by other conditions associated with cognitive

dysfunction have also been known to experience many of these barriers as indicated by Web design guidelines [47].

However, included barriers were related to neurocognitive dysfunction—impaired attention, processing and responding to information slowly and problem-solving—and none were associated with sociocognitive deficits—impaired affect regulation and difficulty processing emotional cues. This is possibly due to affective measurements being overlooked by researchers of included studies.

Barriers were predominantly identified using qualitative research methods and to a lesser extent mixed and quantitative methods. Identified barriers were often not well stated—not including details about the particular user category affected, disability type, hindered activity or task, and how it is hindered—and there was no indication of how restrictive barriers were or how often particular groups of participants encountered them. This can contribute to the development of tentative and inconclusive recommendations that may not be helpful.

Recommended Facilitation Measures

Studies recommended facilitation measures that contribute towards ensuring the use of intuitive navigation, correctly functioning features, simple language, explicit, consistent and easy-to-detect website components, organized content, a flat hierarchical content structure, multimedia formats, and easy-to-operate functions. Facilitation measures recommended by included guidelines focused on improvement strategies that ensure websites are sufficiently operable, understandable, perceivable, and robust. Given the overlap in barriers, it was correctly anticipated that identified facilitation measures would also be mostly in agreement with recommendations for other conditions associated with cognitive deficits.

Facilitation measures were largely developed based on the opinion of researchers conducting studies and consensus among members of international working groups of experts in the area of accessibility. Some researchers [38] disagree with this approach because it does not involve empirical research with people affected by the particular condition when finding ways to meet their needs. However, facilitation measures derived from empirical work were similar to those based on the opinion of researchers conducting studies. Nonetheless, as shown in *Results* section, more focus was placed in different areas for 2 of the 3 principles under which facilitation measures were categorized. Facilitation measures recommended by included guidelines addressed problem areas, whereas other facilitation measures targeted specific barriers.

Facilitation measures were seldom linked to barriers. For example, no facilitation measures recommended by included guidelines had barriers associated with them. Consequently, many facilitation measures were recommended without validation and in a way that makes future validation difficult. This poses a challenge when selecting facilitation measures to address a particular barrier and attempting to increase the effectiveness of a particular facilitation measure.

Coverage of Mental Disorders

As schizophrenia is associated with more severe cognitive deficits than other conditions [48,49] and many participants were also recruited from institutional settings, it was foreseeable that most studies in the area would involve people affected by these 2 conditions. Good cognitive ability is very important when using the Web [38], and the deficits associated with these conditions can put this population at high risk of encountering barriers when using digital technologies such as the Web. Although people affected by depression, anxiety, and bipolar disorder are believed to experience less severe cognitive deficits than those affected by schizophrenia [15], these conditions received similar coverage by included studies. This is possibly due to these conditions being common and the debilitating impact they could still have on the lives of people affected.

Coverage of Digital Technologies

The overwhelming focus on websites out of many digital technologies demonstrates the heavy importance placed on the Web for its usefulness for PwMD. It also acknowledges that there is a need to further optimize Web-based resources. A single 1998 study [35] did not focus on websites but on a multimedia application. This is not surprising as the Web was not widely adopted during that time, but such applications were common.

Types and Suitability of Study Designs

Qualitative methods were suitably adopted for most included studies because they sought to describe and explore technology usage and design for PwMD. The 3 other studies [37-39] investigated the effectiveness of design elements for PwMD and appropriately used quantitative usability testing methods.

It is acknowledged that more granular analysis and reporting of results by mental disorders in studies that involved people with more than 1 MD could potentially reveal a slightly different result. All studies except 3 [37,39-41] noted the classification of MD used when recruiting participants, and this makes it challenging to perform comparisons between results of similar studies and mental disorders and to confidently link results to classifications.

Included studies raise concerns about a bias toward Western culture owing to an absence of research conducted with participants from other cultures. Multicountry studies (eg, [50,51]) have established that culture helps shape technology usage to a great extent.

Participants in included studies ranged widely in age from 18 to over 75 years, and the experiences between younger and older participants were rarely compared or separated. It is important to account for age because it plays a significant role in determining the types of barriers individuals experience when using technology [52,53].

Recency of Research

Findings show that more accessibility and usability research involving PwMD have been done in the last 5 years (10) compared with previous times (3). Considerably more research was done during the same period as revealed by a keyword search of several databases (ie, MEDLINE, PsycARTICLES,

CINAHL, Library, Information Science and Technology Abstracts, Computers and Applied Sciences Complete, and ACM Digital Library) for Web or information and communications technology or digital accessibility or usability and visual (139), mobility (64), cognitive and learning (34), and auditory (17) impairments. This suggests that activity in the area is increasing but not at a rate comparable to similar research done with other populations. Included accessibility guidelines were dated. However, version 2.0 updates for Authoring Tool Accessibility Guidelines and User Agent Accessibility Guidelines are almost stable and referenceable versions that will likely be W3C Recommendations and new Web standards [54,55].

Limitations

Although the literature search was conducted in many databases, results were limited to publications in English. However, no publications were later excluded based on this restriction. Included publications were not limited to those involving empirical work because preliminary searches indicated a paucity of research focusing on the area. As a result, international guidelines were included in the review. However, these guidelines are based on consensus among many experts and not empirical work, which allows for more valid conclusions. Moreover, although identified barriers found in studies were the result of empirical work, not all facilitation measures identified by studies were empirically validated. Nonetheless, as mentioned in *Discussion* section, empirically derived facilitation measures were similar to those based on the expert opinion of researchers conducting studies and were not in conflict with facilitation measures recommended by international guidelines.

Most of the included studies did not use a structured diagnostic classification (eg, ICD or DSM), and this has repercussions for our conclusions being tied to a diagnosis. For instance, it cannot be said unequivocally that persons with a particular diagnosis (eg, depression) experience a certain barrier as reported by those studies that did not use a structured diagnostic classification. Care was also taken to avoid making strong conclusions based on the small number of included studies (13), and it is advised that findings should be interpreted with this in mind.

Implications and Recommendations for Practice and Future Research

Web professionals can now consult a full compilation of research and guidelines–based barriers and facilitation measures relevant to PwMD when developing and optimizing Web-based resources. This will raise awareness of PwMD's needs when using the Web among Web professionals and potentially stimulate further discussion and action within the profession.

The body of research is in need of significant development, and it is too early to make meaningful conclusions on any particular MD, especially based on high-risk symptomatology. For future research, priority should be given to investigating all mental disorders initially. More research in the area is therefore required especially for mood, anxiety, dissociative, somatic, eating, sleep, impulse control, and personality disorders as these have attracted little or no attention.

In agreement with [56-58], an increased effort is needed to investigate the accessibility of technological innovations and health systems. This should be done in a more systematic way with clinically diagnosed samples to obtain conclusive evidence about what barriers exist and how they can be removed. This would involve ensuring each barrier is well stated along with an indication of the level of restriction it causes and frequency of occurrence among the particular user group. Validating strategies targeting the removal of barriers before recommending them as facilitation measures would also be helpful.

Additional actions could be taken by researchers to further develop this area of work. Incorporating valid measures for sociocognitive impairment allows for a more comprehensive evaluation of accessibility for PwMD. It would be important to know if there are cultural differences in the barriers encountered, the level of restriction a particular barrier causes, and/or the frequency of its occurrence. Accessibility studies could also consider a wider range of websites—social networking, e-commerce, education, health—and not just websites targeting PwMD to ensure all aspects of Web usage are investigated.

Conclusions

Indeed, PwMD encounter barriers on the Web, and attempts have been made to remove or reduce these barriers. To the best of our knowledge, these results represent the first attempt to consolidate information on all barriers and facilitation measures investigated for PwMD when using digital technologies in a systematic way. However, it must be taken into consideration that only 13 studies and 3 guidelines meeting the inclusion criteria were identified. These findings also highlight the dire need for more rigorous research to be exhaustive and to have a larger impact on improving the Web for PwMD.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Summary of search concepts and terms.

[PDF File (Adobe PDF File), 45KB - jmir_v18i6e157_app1.pdf]

References

1. Kessler RC, Angermeyer M, Anthony JC, Demyttenaere K, Gasquet I, Gluzman S, et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. World Psychiatry 2007 Oct;6(3):168-176 [FREE Full text] [Medline: 18188442]

2. Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. Lancet 2013 Nov 9;382(9904):1575-1586. [doi: 10.1016/S0140-6736(13)61611-6] [Medline: 23993280]

3. Cieza A, Anczewska M, Ayuso-Mateos JL, Baker M, Bickenbach J, Chatterji S, et al. Understanding the Impact of Brain Disorders: Towards a 'Horizontal Epidemiology' of Psychosocial Difficulties and Their Determinants. PLoS One 2015;10(9):e0136271 [FREE Full text] [doi: 10.1371/journal.pone.0136271] [Medline: 26352911]

4. Navidian A, Zaheden F. Burden experienced by family caregivers of patients with mental disorders. Pakistan Journal of Psychological Research 2008;23(1):19-28.

5. World Health Assembly 65. WHO. 2012. Global burden of mental disorders and the need for a comprehensive, coordinated response from health and social sectors at the country level URL: https://extranet.who.int/iris/restricted/bitstream/10665/80478/1/A65_R4-en.pdf [WebCite Cache ID 6ds4fDQwe]

6. Kalckreuth S, Trefflich F, Rummel-Kluge C. Mental health related Internet use among psychiatric patients: a cross-sectional analysis. BMC Psychiatry 2014;14:368 [FREE Full text] [doi: 10.1186/s12888-014-0368-7] [Medline: 25599722]

7. Chen J, Zhu S. Online Information Searches and Help Seeking for Mental Health Problems in Urban China. Adm Policy Ment Health 2015 May 16:NA Epub ahead of print(forthcoming). [doi: 10.1007/s10488-015-0657-6] [Medline: 25981055]

8. Kummervold PE, Gammon D, Bergvik S, Johnsen JK, Hasvold T, Rosenvinge JH. Social support in a wired world: use of online mental health forums in Norway. Nord J Psychiatry 2002;56(1):59-65. [doi: 10.1080/08039480252803945] [Medline: 11869468]

9. Powell J, McCarthy N, Eysenbach G. Cross-sectional survey of users of Internet depression communities. BMC Psychiatry 2003 Dec 10;3:19 [FREE Full text] [doi: 10.1186/1471-244X-3-19] [Medline: 14664725]

10. Andersson G, Cuijpers P, Carlbring P, Riper H, Hedman E. Guided Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta-analysis. World Psychiatry 2014 Oct;13(3):288-295 [FREE Full text] [doi: 10.1002/wps.20151] [Medline: 25273302]

11. Kuester A, Niemeyer H, Knaevelsrud C. Internet-based interventions for posttraumatic stress: A meta-analysis of randomized controlled trials. Clin Psychol Rev 2016 Feb;43:1-16. [doi: 10.1016/j.cpr.2015.11.004] [Medline: 26655959]

12. Van Voorhees B, Vanderplough-Booth K, Fogel J, Gladstone T, Bell C, Stuart S, et al. Integrative internet-based depression prevention for adolescents: a randomized clinical trial in primary care for vulnerability and protective factors. J Can Acad Child Adolesc Psychiatry 2008 Nov;17(4):184-196 [FREE Full text] [Medline: 19018321]

13. O'Kearney R, Kang K, Christensen H, Griffiths K. A controlled trial of a school-based Internet program for reducing depressive symptoms in adolescent girls. Depress Anxiety 2009;26(1):65-72. [doi: 10.1002/da.20507] [Medline: 18828141]

14. Johnson GM. Functional Internet Literacy: required cognitive skills with implications for instruction. E-Learning 2007;4(4):433-441. [doi: 10.2304/elea.2007.4.4.433]

15. Millan MJ, Agid Y, Brüne M, Bullmore ET, Carter CS, Clayton NS, et al. Cognitive dysfunction in psychiatric disorders: characteristics, causes and the quest for improved therapy. Nat Rev Drug Discov 2012 Feb;11(2):141-168. [doi: 10.1038/nrd3628] [Medline: 22293568]

16. Short J, Williams E, Christie B. The social psychology of telecommunications. London: Wiley; 1976.

17. Lea M, Spears R. Love at first byte? Building personal relationships over computer networks. In: Wood JT, Duck S, editors. Under-studied relationships: off the beaten track. Thousand Oaks, CA: SAGE Publications; 1995.

18. WALTHER JB. Interpersonal Effects in Computer-Mediated Interaction: A Relational Perspective. Communication Research 1992 Feb 01;19(1):52-90. [doi: 10.1177/009365092019001003]

19. WALTHER JB. Computer-Mediated Communication: Impersonal, Interpersonal, and Hyperpersonal Interaction. Communication Research 1996 Feb 01;23(1):3-43. [doi: 10.1177/009365096023001001]

20. Mariger H. Cognitive Disabilities and the Web: Where Accessibility and Usability Meet?. 2006. URL: http://ncdae.org/resources/articles/cognitive/ [accessed 2015-12-18] [WebCite Cache ID 6ds3w1PeD]

21. Hudson R, Weakley R, Firminger P. An Accessibility Frontier: Cognitive disabilities and learning difficulties. 2004. URL: http://usability.com.au/2004/12/an-accessibility-frontier-cognitive-disabilities-and-learning-difficulties-2004/ [accessed 2015-12-18] [WebCite Cache ID 6ds4HcUpp]

22. Bernard R, Sabariego C, Baldwin D, Abou-Zahra S, Cieza A. BETTER-project: web accessibility for persons with mental disorders. In: Kurosu M, editor. Human-Computer Interaction, Users and Contexts 17th International Conference, HCI International. Switzerland: Springer International Publishing; 2015:25-34.

23. World Wide Web Consortium. 2008. Web Content Accessibility Guidelines (WCAG) 2.0 URL: http://www.w3.org/TR/WCAG20/ [accessed 2015-12-18] [WebCite Cache ID 6ds3MFPKT]

24. Denison HJ, Dodds RM, Ntani G, Cooper R, Cooper C, Sayer AA, et al. How to get started with a systematic review in epidemiology: an introductory guide for early career researchers. Arch Public Health 2013;71(1):21 [FREE Full text] [doi: 10.1186/0778-7367-71-21] [Medline: 23919540]

25. World Health Organization. International classification of functioning, disability and health: ICF. Geneva: World Health Organization; 2001.

26. World Wide Web Consortium. Introduction to Understanding WCAG 2 URL: https://www.w3.org/TR/UNDERSTANDING-WCAG20/intro.html [accessed 2016-01-27] [WebCite Cache ID 6eqZRHcpt]

27. World Wide Web Consortium. WCAG 2 at a Glance. 2012. URL: https://www.w3.org/WAI/WCAG20/glance/ [accessed 2016-03-03] [WebCite Cache ID 6fj9a8ph6]

28. World Wide Web Consortium. 2014. ATAG at a Glance URL: https://www.w3.org/WAI/intro/atag-glance.php [accessed 2016-03-03] [WebCite Cache ID 6fjAN1sn8]

29. World Wide Web Consortium. 2005. User Agent Accessibility Guidelines (UAAG) Overview URL: http://www.w3.org/WAI/intro/uaag.php [accessed 2015-12-18] [WebCite Cache ID 6ds9GQNvX]

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

31. Bulger J. Jenniferbulger. 2002. A usability study of mental health websites with an emphasis on homepage design: Performance and preferences of those with anxiety disorders URL: http://www.jenniferbulger.com/usability/Master/'s_paper.pdf [accessed 2015-12-18] [WebCite Cache ID 6ds5Gj8q7]

32. Brunette MF, Ferron JC, Devitt T, Geiger P, Martin WM, Pratt S, et al. Do smoking cessation websites meet the needs of smokers with severe mental illnesses? Health Educ Res 2012 Apr;27(2):183-190 [FREE Full text] [doi: 10.1093/her/cyr092] [Medline: 21987478]

33. Ferron JC, Brunette MF, McHugo GJ, Devitt TS, Martin WM, Drake RE. Developing a quit smoking website that is usable by people with severe mental illnesses. Psychiatr Rehabil J 2011;35(2):111-116. [doi: 10.2975/35.2.2011.111.116] [Medline: 22020840]

34. Good A, Sambhanthan A. Accessing web based health care and resources for mental health: interface design considerations for people experiencing mental illness. In: Marcus A, editor. Accessing web based health careresources for mental health: interface design considerations for people experiencing mental illness. Switzerland: Springer International Publishing; 2014:978-973.

35. Jimison HB, Sher PP, Appleyard R, LeVernois Y. The use of multimedia in the informed consent process. J Am Med Inform Assoc 1998;5(3):245-256 [FREE Full text] [Medline: 9609494]

36. Johansson S, Gulliksen J, Lantz A. Cognitive Accessibility for Mentally Disabled Persons. In: Abascal J, Barbosa S, Fetter M, Gross T, Palanque P, Winckler M, editors. Human-Computer Interaction – INTERACT 2015. Switzerland: Springer International Publishing; 2015:418-435.

37. Rotondi AJ, Eack SM, Hanusa BH, Spring MB, Haas GL. Critical design elements of e-health applications for users with severe mental illness: singular focus, simple architecture, prominent contents, explicit navigation, and inclusive hyperlinks. Schizophr Bull 2015 Mar;41(2):440-448. [doi: 10.1093/schbul/sbt194] [Medline: 24375458]

38. Rotondi AJ, Sinkule J, Haas GL, Spring MB, Litschge 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. Psychol Serv 2007 Aug;4(3):202-224 [FREE Full text] [doi: 10.1037/1541-1559.4.3.202] [Medline: 26321884]

39. Rotondi AJ, Spring M, Grady J, Simpson R, Luther J, Abebe KZ, et al. Use of a Fractional Factorial Experiment to Assess the E-healthcare Application Design Needs of Persons With Dual Diagnosis. Journal of Dual Diagnosis 2012 Nov;8(4):277-282. [doi: 10.1080/15504263.2012.723454]

40. Todd NJ, Jones SH, Lobban FA. What do service users with bipolar disorder want from a web-based self-management intervention? A qualitative focus group study. Clin Psychol Psychother 2013;20(6):531-543. [doi: 10.1002/cpp.1804] [Medline: 22715161]

41. Schrank B, Sibitz I, Unger A, Amering M. How patients with schizophrenia use the internet: qualitative study. J Med Internet Res 2010;12(5):e70 [FREE Full text] [doi: 10.2196/jmir.1550] [Medline: 21169176]

42. Black AC, Serowik KL, Schensul JJ, Bowen AM, Rosen MI. Build a better mouse: directly-observed issues in computer use for adults with SMI. Psychiatr Q 2013 Mar;84(1):81-92 [FREE Full text] [doi: 10.1007/s11126-012-9229-z] [Medline: 22711454]

43. Treviranus J, McCathieNevile C, Jacobs I, Richards J. World Wide Web Consortium. 2000. Authoring Tool Accessibility Guidelines 1.0 URL: http://www.w3.org/TR/2000/REC-ATAG10-20000203/ [accessed 2015-12-18] [WebCite Cache ID 6ds7bYajq]

44. Jacobs I, Gunderson J, Hansen E. World Wide Web Consortium. 2002. User Agent Accessibility Guidelines 1.0 URL: http://www.w3.org/TR/WAI-USERAGENT/ [accessed 2015-12-18] [WebCite Cache ID 6ds88jvoR]

45. World Wide Web Consortium. 2008. How WAI Develops Accessibility Guidelines through the W3C Process: Milestones and Opportunities to Contribute URL: http://www.w3.org/WAI/intro/w3c-process.php [accessed 2015-12-18] [WebCite Cache ID 6ds8WQ8Zl]

46. World Wide Web Consortium. 2012. Diversity of Web Users: How People with Disabilities Use the Web URL: http://www.w3.org/WAI/intro/people-use-web/diversity [accessed 2015-12-18] [WebCite Cache ID 6ds8eptA6]

47. Friedman M, Bryen D. Web accessibility design recommendations for people with cognitive disabilities. Technology & Disability 2007;19(4):205-212.

48. Shrivastava AK, Johnston ME. Cognitive neurosciences: A new paradigm in management and outcome of schizophrenia. Indian J Psychiatry 2010 Apr;52(2):100-105 [FREE Full text] [doi: 10.4103/0019-5545.64575] [Medline: 20838495]

49. Bora E, Yucel M, Pantelis C. Cognitive functioning in schizophrenia, schizoaffective disorder and affective psychoses: meta-analytic study. Br J Psychiatry 2009 Dec;195(6):475-482 [FREE Full text] [doi: 10.1192/bjp.bp.108.055731] [Medline: 19949193]

50. Calhoun KJ, Teng JTC, Cheon MJ. Impact of national culture on information technology usage behaviour: An exploratory study of decision making in Korea and the USA. Behaviour & Information Technology 2010 Nov 08;21(4):293-302. [doi: 10.1080/0144929021000013491]

51. Park C, Jun J. A cross‑cultural comparison of Internet buying behavior. International Marketing Review 2003 Oct;20(5):534-553. [doi: 10.1108/02651330310498771]

52. Díaz-Bossini J, Moreno L. Accessibility to Mobile Interfaces for Older People. Procedia Computer Science 2014;27:57-66. [doi: 10.1016/j.procs.2014.02.008]

53. de Lara SMA, Fortes RPDM, Russo CM, Freire AP. A study on the acceptance of website interaction aids by older adults. Univ Access Inf Soc 2015 Jun 24:1-16. [doi: 10.1007/s10209-015-0419-y]

54. World Wide Web Consortium. 2013. Authoring Tool Accessibility Guidelines (ATAG) Overview URL: http://www.w3.org/WAI/intro/atag.php [accessed 2015-12-18] [WebCite Cache ID 6ds9C55we]

55. World Wide Web Consortium. 2005. User Agent Accessibility Guidelines (UAAG) Overview URL: http://www.w3.org/WAI/intro/uaag.php [accessed 2015-12-18] [WebCite Cache ID 6ds9GQNvX]

56. Davies EB, Morriss R, Glazebrook C. Computer-delivered and web-based interventions to improve depression, anxiety, and psychological well-being of university students: a systematic review and meta-analysis. J Med Internet Res 2014;16(5):e130 [FREE Full text] [doi: 10.2196/jmir.3142] [Medline: 24836465]

57. 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]

58. Renton T, Tang H, Ennis N, Cusimano MD, Bhalerao S, Schweizer TA, et al. Web-based intervention programs for depression: a scoping review and evaluation. J Med Internet Res 2014;16(9):e209 [FREE Full text] [doi: 10.2196/jmir.3147] [Medline: 25249003]

Abbreviations

MD: mental disorders
PwMD: people with mental disorders
SMI: severe mental illness
USA: United States of America
W3C: World Wide Web Consortium

XSL·FO
RenderX

XSL•FO

RenderX

Original Paper

Do Web-based Mental Health Literacy Interventions Improve the Mental Health Literacy of Adult Consumers? Results From a Systematic Review

Bianca Brijnath[1*], B.A. (Hons), PhD; Joanne Protheroe[2*], MBChB, PhD, FRCGP; Kamal Ram Mahtani[3*], BSc, PhD, MBBS, PGDip, MRCGP; Josefine Antoniades[4], BSc (Hons)

[1]Curtin University, School of Occupational Therapy and Social Work, Perth, Australia

[2]Institute of Primary Care and Health Sciences, Arthritis Research UK Primary Care Centre, Keele University, Oxford, United Kingdom

[3]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom

[4]Department of General Practice, Faculty of Medicine, Nursing and Health Sciences, Monash University, Notting Hill, Australia

[*]these authors contributed equally

Corresponding Author:
Bianca Brijnath, B.A. (Hons), PhD
Curtin University
School of Occupational Therapy and Social Work
Building 401
Bentley Campus
Perth, 6152
Australia
Phone: 61 8 9266 5206
Fax: 61 8 9266 3636
Email: bianca.brijnath@curtin.edu.au

Abstract

Background: Low levels of mental health literacy (MHL) have been identified as an important contributor to the mental health treatment gap. Interventions to improve MHL have used traditional media (eg, community talks, print media) and new platforms (eg, the Internet). Evaluations of interventions using conventional media show improvements in MHL improve community recognition of mental illness as well as knowledge, attitude, and intended behaviors toward people having mental illness. However, the potential of new media, such as the Internet, to enhance MHL has yet to be systematically evaluated.

Objective: Study aims were twofold: (1) To systematically appraise the efficacy of Web-based interventions in improving MHL. (2) To establish if increases in MHL translated into improvement in individual health seeking and health outcomes as well as reductions in stigma toward people with mental illness.

Methods: We conducted a systematic search and appraisal of all original research published between 2000 and 2015 that evaluated Web-based interventions to improve MHL. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines were used to report findings.

Results: Fourteen studies were included: 10 randomized controlled trials and 4 quasi-experimental studies. Seven studies were conducted in Australia. A variety of Web-based interventions were identified ranging from linear, static websites to highly interactive interventions such as social media games. Some Web-based interventions were specifically designed for people living with mental illness whereas others were applicable to the general population. Interventions were more likely to be successful if they included "active ingredients" such as a structured program, were tailored to specific populations, delivered evidenced-based content, and promoted interactivity and experiential learning.

Conclusions: Web-based interventions targeting MHL are more likely to be successful if they include active ingredients. Improvements in MHL see concomitant improvements in health outcomes, especially for individuals with mild to moderate depression. The most promising interventions suited to this cohort appear to be MoodGYM and BluePages, 2 interventions from Australia. However, the relationship between MHL and formal and informal help seeking is less clear; self-stigma appears to be an important mediator with results showing that despite improvements in MHL and community attitudes to mental illness, individuals with mental illness still seek help at relatively low rates. Overall, the Internet is a viable method to improve MHL.

Future studies could explore how new technology interfaces (eg, mobile phones vs computers) can help improve MHL, mental health outcomes, and reduce stigma.

Introduction

Despite the high global prevalence of mental illness [1], a significant treatment gap remains between those requiring care and those receiving care. In high-income English-speaking countries, such as the United States, the United Kingdom, and Australia, the prevalence of mental illness ranges from 14.9% to 24.6%, yet the treatment gap is 40%-65% [2-4]. One of the main reasons for this gap is low levels of mental health literacy (MHL) [5,6]. Defined as "the knowledge and beliefs about mental disorders, which aids their recognition, management or prevention" [7], MHL consists of 6 components: (1) the ability to recognize mental illnesses; (2) knowledge and beliefs about risk factors and causes; (3) knowledge about self-help interventions; (4) knowledge and beliefs about professional help available; (5) attitudes that facilitate recognition and appropriate help seeking; and (6) knowledge about how to seek appropriate mental health information [8].

Several studies have now conclusively shown that improvements in MHL improve community recognition of mental illness as well as knowledge, attitude, and intended behaviors toward people having mental illness [6,9-13]. The relationship between MHL and reduction in stigma toward people living with mental illness is still unclear [13]. However, these results have mainly been derived from evaluations of large-scale community mental health awareness campaigns delivered through traditional media such as television, radio, and print media; interpersonal contact with a person with a mental illness; and public seminars and community talks [6,9-13]. The potential of new media, such as the Internet, to enhance MHL has also been explored, but the actual effect on increasing MHL has yet to be systematically evaluated. Given the high rate of Internet penetration in the general population—more than 80% in the developed world—the Internet is an ideal medium through which to reach significant numbers of people at relatively low cost [14].

Currently, many MHL interventions are embedded within other interventions and it is difficult to disentangle the efficacy of each component of the intervention from the others (for review of interventions see [6,15]). However, identifying the efficacy of Web-based MHL interventions is important to enable a more strategic scale-up of what actually works and to increase the cost-effectiveness of any future intervention by discarding unsuccessful elements. A rigorous evidence-based cost-effective MHL intervention if successfully delivered via the web can harness the potential of the Internet, thereby increasing MHL at the population level and potentially delivering significant improvements in mental health outcomes among those living with mental illness.

To facilitate the development of such an intervention, the aims of this systematic review were twofold: (1) to collate the existing evidence to establish the efficacy of Web-based interventions that seek to improve MHL and (2) to establish, where possible, whether improvements in MHL translate into improvements in individual health seeking, reductions in stigmatizing attitudes toward people living with mental illness, and better health outcomes for individuals living with mental illness.

Methods

Search Strategy

This review was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [16]. To identify eligible studies 6 databases were searched: PsycINFO, EMBASE, PubMed, CINAHL, and Web of Science. The search was conducted from April to August 2015 and search results were limited to English language, peer-reviewed articles published between 2000 and 2015. We did not search for articles published before 2000 because the global Internet penetration was only 6.5% at that time and we did not envisage any Web-based interventions targeting MHL before 2000. Key terms to identify studies included the following: Mental health literacy (Mental health, mental illness*, mental disorder*, mental disease*, depression AND literacy/*health literacy) and Internet (internet* or online* or web or World Wide Web or social media or website or surfing). Articles identified by the database search were screened to assess relevance to the aims. In addition, Google Scholar and selected reference lists were also searched to identify additional studies of interest. This review is a registered PROSPERO review: CRD42015025572.

Study Inclusion and Exclusion Criteria

To be included, studies were required to meet the following criteria: (1) published in English in a peer-reviewed journal, (2) described a Web-based intervention with either the primary or secondary aim to improve MHL and included a measure of MHL or a component thereof such as mental health knowledge, (3) reported original research of a quantitative, qualitative, or mixed methods design, (4) included participants who were 17 years or older, and (5) included community members, family members and carers, and/or patients. Publications were excluded if participants in the study were health care professionals, the publications comprised commentaries or editorials rather than empirical research, or were published before 2000.

Selection of Studies

Two authors (BB and JA) independently assessed relevant titles and abstracts. Selected studies were obtained in full text and reviewed in detail by 3 authors (BB, JA, and JP). Where any

XSL•FO

RenderX

discrepancies arose, a consensus was reached through discussion or, if necessary, referral to another author (KM). After full-text review a number of studies were excluded as being irrelevant and the final number of included studies was obtained.

Data Extraction

Data extraction was carried out independently by 3 authors (BB, JA, and JP). Where any discrepancies arose they were referred to a fourth author (KM). Data were extracted onto a predesigned worksheet relevant to our outcomes of interest.

Quality Assessment

To ensure methodological rigor in the review process, all included studies were appraised by a minimum of 2 authors (JA, BB, or JP) for quality in accordance with the United Kingdom's National Institute for Health and Care Excellence (NICE) guidelines and methodology [17]. As it is difficult to blind participants for behavioral treatment, we redefined the criterion regarding the blinding of participants. If blinding was not feasible, item 4 of the quality assessment was scored positive (+) if the credibility of the treatments was evaluated and treatments were equally credible and acceptable to participants;

that is, control as well as intervention could be perceived to be an intervention in its own right [18].

Results

Through the literature search, 571 potential records were identified (Figure 1); however, after the removal of duplicates, 448 studies were included for review based on title and abstract alone. Of the 448 studies, 26 were retained for full-text review. Full-text articles were reviewed by a minimum of 2 reviewers (BB, JA, JP) and were assessed for suitability for inclusion in accordance with the inclusion and exclusion criteria. During this process a further 12 papers were excluded as they did not meet the inclusion criteria of this review [19-30] (see Figure 1 – PRISMA flowchart for reasons). Therefore 14 articles were retained for inclusion [31-44]. Of these 14 papers, 2 papers reported on the same large randomized controlled trial (RCT) [31,36] but reported on different outcomes and were included as separate papers. However, this has been taken into consideration in the analysis for this review. The interrater agreement of the quality assessment was 84% and any disagreement between assessments after full-text review was resolved through consensus.

Figure 1. PRISMA flowchart.
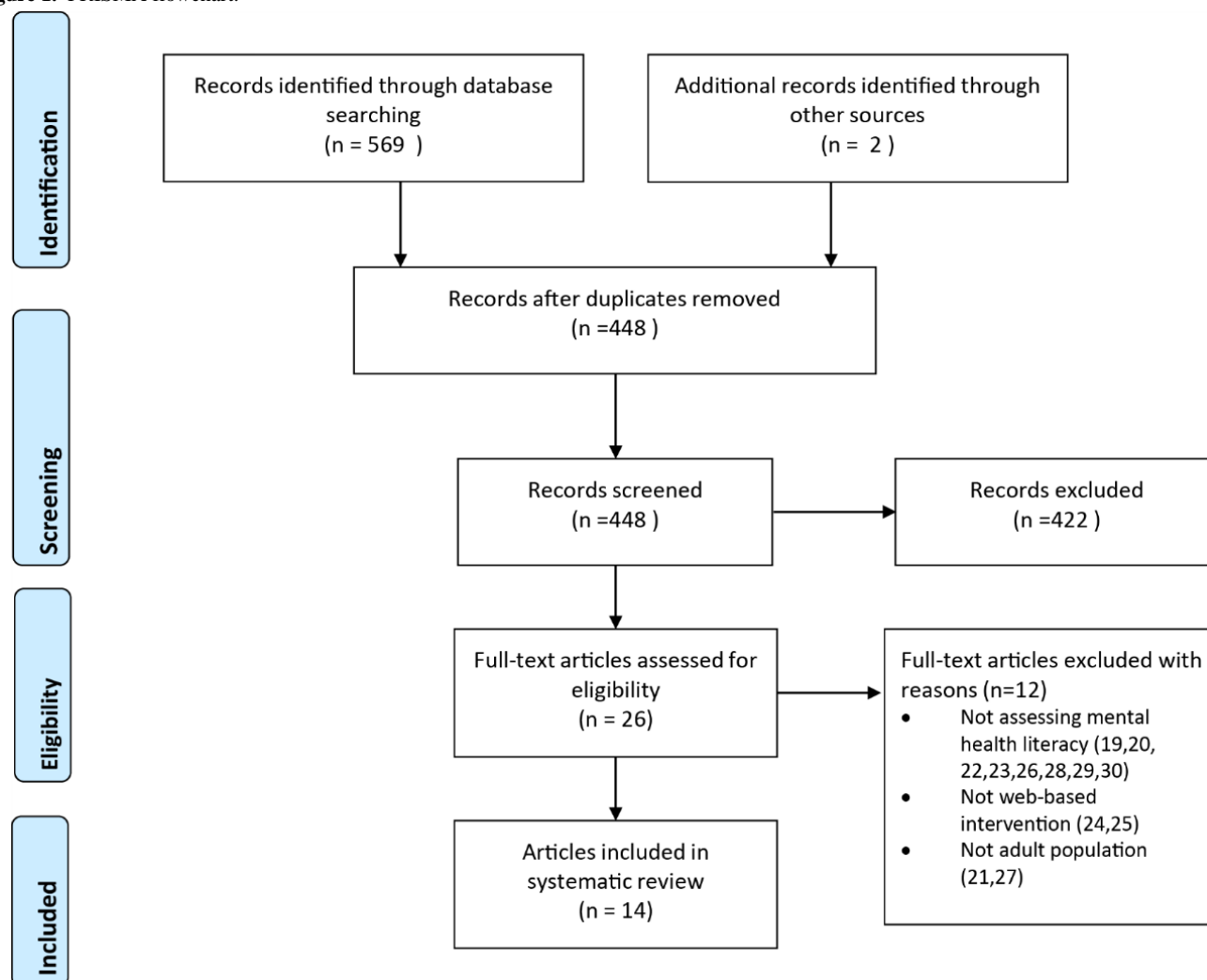
Types of Studies

Of the 14 included studies, 10 were RCTs [31-34,36-38,40,41,44] and 4 were repeated-measures studies [35,39,42,43]. Two articles reported on the same trial [31,36], 7 studies were conducted in Australia [31,32,34,36-38,43,44], 4 studies were from the United States [33,35,41,42], 1 study was conducted in Hong Kong [39], and finally 1 study was Norwegian [40] but was reporting a trial of Australian self-help interventions, MoodGYM and BluePages translated into Norwegian. None of the reviewed studies included a qualitative exploration of the effect of the intervention on MHL, health seeking, stigma, or health outcomes. Five of the studies were complex interventions comprising 2 or more components [31,34,36,37,40].

Participant Characteristics

Across the 14 studies the total pool of participants was 2605 individuals. Most studies included adult participants with clinical indication of a mental illness [32,33,35,37-39,42-44], and only 4 studies (5 papers) specifically recruited participants with mental health problems [31,34,36,40,41]. Two studies [33,42] specifically focused on family members and carers, 7 on the general community [32,35,37-39,43,44], and 1 study had a combined focus on patients and carers [41] (Multimedia Appendix 1). Despite the heterogeneity of target populations, comparability within and across groups was possible because most used the same constructs and measures; 7 studies used the Depression Literacy Questionnaire (D-Lit) alone or in combination with others to measure MHL [31,34-38,44]; 4 of the 6 studies reporting on stigma used the Depression Stigma Scale (DSS) [34,37,38,44] alone or in combination with other scales; 3 of the 5 studies reporting on help seeking used the General Help-Seeking Questionnaire (GHSQ) [32,37,44]; and

10 studies that included a measure of mental illness symptomatology used the Center for Epidemiologic Studies-Depression (CES-D) [31,32,34-36,38,40,41,43,44]. Further information is detailed in Multimedia Appendix 1: Study overview and characteristics. Eleven studies reported unequal gender representation with an average of 67.9% females [31,32,35-37,39-44].

Study Quality Indicators

A summary of risk of bias and quality indicators for RCTs can be found in Figure 2 and Table 1, respectively, and Table 2 for quality indicators and risk of bias in non-RCTs. Using the NICE guidelines to assess study quality, we found that a substantial number of studies reported high attrition rates (>20%) [31,33,34,37,39,40,43]; however, most studies included robust means of handling missing data such as intention-to-treat (ITT) analysis, which renders a conservative estimate of intervention effects [45]. Furthermore, although often practically unavoidable, some studies had small sample sizes, frequently related to recruitment and/or retention difficulties, and others recruited from limited pools of participants such as social clubs, student populations, or organizations, which may limit the generalizability of the findings (Tables 1 and 2). While blinding of participants is problematic for this type of intervention, which we have taken into consideration, most studies did not blind investigators who were involved with assessing the data, which could introduce detection bias. Moreover, in included RCTs randomization procedures were not consistently reported (Figure 2) and although the results of many of the included studies were encouraging, in some cases data were only collected immediately before and after the intervention with no subsequent follow-up, hence the sustainability of the interventions remains unclear (Table 1).

Table 1. Quality indicators and limitations for randomized controlled trials.

Authors	Recruitment	Data collection	Attrition	Adherence	Limitations
Christensen et al [31]	Election roll	Pre/post	Lost to follow-up: 18% for BluePages, 33% for MoodGYM, and 12% for control	Yes	Attrition rate Longer follow-up desirable
Griffiths et al [36]	Election roll	Pre/post	Lost to follow-up: 18% for BluePages, 33% for MoodGYM, and 12% for control	Not reported, but reported in Christiansen et al [31]	Small effect sizes Attrition rates Longer follow-up desirable
Costin et al [32]	Election roll	Pre/post (3 weeks after intervention)	Control (high/low distress): 14.5% Intervention (basic): 15.3% Intervention (enhanced): 17%	Yes	Power calculations suggest larger sample required No follow-up
Kiropoulos et al [38]	Welfare and social groups	Pre/post/1 week	0% (one-off access to website)	Not applicable	Sample may not be representative Researcher present during intervention Longer follow-up desirable
Rotondi et al [41]	Community mental health centers inpatient units	Pre/post/3, 6, 12 months	Patients: 3% Carers: 17%	Yes, high adherence	Small sample size Face-to-face workshop before intervention
Taylor-Rodgers and Batterham [44]	University	Pre/post	Control: 18% Intervention: 15%	Yes, 65.4% of intervention and 70.4% of control viewed all 3 Web pages	Small sample size University-based sample Longer follow-up desirable
Lintvedt et al [40]	University	Pre/post/2 months	Control: 28% Intervention: 46.9%	Not reported	Attrition rate University sample Longer follow-up desirable
Deitz et al [33]	Employees in 1 worksite	Pre/post	Not adequately reported: given response rate for intervention: 96%, control: 98%, but 22% of total sample did not view Web-based material	Not reported	Could not monitor "dosage" of intervention Limited sample Longer follow-up desirable
Farrer et al [34]	Mental health support hotline (Lifeline)	Pre/post/6-12 months	31% at postintervention 41% at 6-month follow-up	Not reported	Small sample size Attrition Adherence not reported
Gulliver et al [37]	Sports organizations	Pre/intervention week 1-2/post/3-6 months	49.2% at follow-up	Not reported	Small sample Study underpowered Attrition

XSL•FO
RenderX

Table 2. Quality indicators for nonrandomized controlled trials.

Authors	Recruitment	Random-ization	Blinding[a]	Data collection	Attrition	Missing data handling	Adherence	Limitations
Shandley et al [43]	Not reported	N/A[b]	N/A	Pre/post/2 months	Post: 42.1% Follow-up: 62.4%	ITT[c]	Yes	Attrition; Limited adherence
Finkelstein and Lapshin [35]	University medical school	N/A	N/A	Pre-post (immediate)	Not applicable/data collected immediate pre/post	N/A	N/A	Follow-up; University sample
Li et al [39]	University	N/A	N/A	Pre/post	Post: 42.1%	ITT	Not reported in detail	Attrition; Small, university sample
Roy et al [42]	Military services	N/A	N/A	Pre/post/optional at 10 days	Post: 0% (only 1 event of using website) Optional follow-up: (74.4%)	N/A	Not reported	Lack of reporting on methods; Short, optional follow-up

[a] Blinding of participants and/or personnel.

[b] N/A: not applicable.

[c] ITT: intention-to-treat.

Figure 2. Risk of bias for randomized controlled trials.

Type of bias — Authors	1. Adequate sequence generation	2. Adequate allocation concealment	3. Blinding patients	4. Blinding assessors	5. Incomplete data addressed (Attrition)	6. incomplete data addressed (ITT analysis)	7. Similarity at baseline	8. Overall risk of bias
Christensen et al [31]	++	++	+	-	++	++	++	++
Griffiths et al [36]	++	++	+	-	++	++	++	++
Costin et al [32]	++	-	+	-	++	+	++	+
Kiropoulos et al [38]	+	-	+	-	NA	NA	++	++
Rotondi et al [41]	+	+	-	-	++	NA	++	+
Taylor-Rodgers et al [44]	++	+	+	-	++	++	++	++
Lintvedt et al [40]	++	++	+	-	-	++	++	++
Deitz et al [33]	-	+	-	-	+	-	++	+
Farrer et al [34]	+	++	+	-	+	++	+	+
Gulliver et al [37]	++	++	-	-	-	++	++	+

Low risk of bias	++ (green)	High risk of bias	- (red)	Unclear risk of bias	+ (yellow)	Not Applicable	NA (blue)

Impact on Mental Health Literacy

Across the reviewed studies there were an assortment of Web-based interventions (Multimedia Appendix 1) that targeted MHL as a primary outcome. Five of these studies employed samples with no prerequisite of symptomatology of mental illness [33,38,39,42,44].

In an innovative study, Li et al [39] tested a social network game, " Ching Ching Story," specifically designed to improve knowledge about mental health problems and the results were encouraging with significant improvements in MHL (Multimedia Appendix 2). Also targeting young adults, Taylor-Rodgers and Batterham [44] assessed the efficacy of a 3-week psychoeducational intervention based on vignettes about mental health problems on MHL as well as stigma, and help-seeking attitudes and intentions (Multimedia Appendix 1). Results suggested that the intervention was moderately effective (d=0.65) in improving anxiety literacy but not depression or suicide literacy and that there was a moderate change (d=0.58) in help-seeking attitudes, in particular toward seeking help from primary care providers (d=0.53; Multimedia Appendix 2).

Targeting carers of children between the ages of 5 and 21 years, Deitz et al [33] reported significant increases in overall knowledge of mental health problems using a 32-item questionnaire on the knowledge of childhood depression and anxiety (P=.008) and improved self-efficacy using a 9-item questionnaire on treatment seeking self-efficacy in handling mental health problems in children (P=.001; Multimedia Appendix 2). These changes resulted from a narrated and interactive Web-based mental health program. However, the program created no change in attitudes to help seeking or toward mental health problems (Multimedia Appendix 2). Roy et al [42] reported improved posttraumatic stress disorder (PTSD) knowledge, as measured using a 25-item PTSD knowledge questionnaire, at postintervention assessment after the use of an educational website for PTSD for the families of military service members specifically designed to increase PTSD knowledge and thereby support for returned military personnel. However, the duration of the intervention was unclear and significant attrition at follow-up was reported (74.4%; Multimedia Appendix 1 and Table 2, respectively). Similarly, the results of an RCT of 3 Web-based interventions (Multimedia Appendix 1) aimed at improving help seeking in young athletes

XSL•FO
RenderX

indicated significant improvements in depression and anxiety literacy levels (Hedges' g=0.90 and 0.90, respectively) compared with all other conditions [37] (Multimedia Appendix 2).

Addressing an extensive gap in the literature, Kiropoulos et al [38] evaluated an Internet-based, multilingual depression information resource targeted at Greek and Italian migrants. The results were encouraging with significant improvements in depression literacy and personal stigma (Multimedia Appendix 2); however, as in other studies, the sustainability of the intervention needs further exploration because participants were only followed up 1 week after the intervention (Table 1).

Although MHL was not the primary aim of the intervention, Shandley et al [43] evaluated a Web-based, CBT-based gaming intervention "Reach Out Central" aimed at supporting mental health in young adults, in particular targeting males (Multimedia Appendix 1). Outcomes suggested significant increases in help-seeking willingness (η^2=.06), particularly for women, and slight improvements in MHL, but only for female participants (Multimedia Appendix 2).

In an RCT testing personalized eHealth cards (Multimedia Appendix 1) to improve help seeking and MHL, no significant results were reported on help seeking or MHL measures. A higher, but nonsignificant, number of positive beliefs about formal help sources and therapy for depression were recorded in the intervention arm (Multimedia Appendix 2). On the other hand, Finkelstein and Lapshin [35] found that their interactive, Web-based educational intervention for depression stigma was not only effective in improving depression stigma, but also significantly increased depression literacy (through the assessment of knowledge and resistance to treatment; Multimedia Appendices 1 and 2).

Three studies investigated the effect of Web-based depression interventions on MHL in populations with elevated depressive symptoms [31,34,40] (Multimedia Appendix 1). Christensen et al [31] conducted a large-scale RCT investigating the effect of BluePages, a depression literacy website, and MoodGYM, a Web-based CBT intervention. Participants in both interventions were followed up on a weekly basis by the research team, providing measurements on depression symptomology, dysfunctional thoughts, and CBT literacy. As hypothesized, both interventions were effective in improving depression literacy relative to the control group. The depression literacy intervention was most effective compared with the CBT intervention and control arm in improving depression literacy; similarly, the CBT intervention was most efficacious in improving CBT literacy (Multimedia Appendix 2).

Lintvedt et al [40] also assessed the effectiveness of BluePages and MoodGYM in Norwegian in improving MHL around depression and CBT in a sample of Norwegian university students. However, in this instance there was no follow-up of participants. Participants were assigned to either the intervention condition, which included access to both self-help websites, or a control condition (waitlist). Results further support the efficacy of MoodGYM and BluePages; the intervention significantly improved depression and CBT literacy and decreased depressive symptoms across all outcome measures, even without the weekly

tracking previously reported by Christensen et al [31] (Multimedia Appendix 2). In an Australian study of individuals with psychological distress a comparable paradigm was employed [34]; participants were allocated to a combination of MoodGYM (6 weeks) and BluePages (1 week) without tracking (weekly 10-minute counselor phone call), tracking only, or control condition (Multimedia Appendix 1). Although CBT literacy significantly improved in Web-intervention conditions (d=0.71 and 0.80 without and with tracking, respectively), overall the intervention did not render a significant improvement in depression literacy and stigma. There did appear to be a short-term improvement in depression literacy and stigma in Web-based conditions, but this improvement was not sustained at 12-month follow-up (Multimedia Appendix 2). As suggested by the authors, these results suggest a dose-dependent effect of the psychoeducational intervention (BluePages) given the success in other trials in which the exposure to intervention content was of a more substantial duration [31,40].

Impact on Seeking Help for Mental Illness

One study reported positive outcomes for help-seeking behaviors [42] after Web-based interventions [42]. While not reporting details relating to the method of data collection, 57% of carers who returned to a PTSD psychoeducational website 10 days after intervention reported having taken action in facilitating help for family member with suspected PTSD, including discussing symptoms and encouraging family member to seek help [42] (Multimedia Appendix 2).

Conversely, 2 studies found no improvement in formal or informal help seeking: Costin et al [32] found no indication that eHealth cards improved help-seeking intentions or actual help seeking among young people; neither did a Web-based mental health program for parents improve attitudes toward help seeking [33] (Multimedia Appendix 2). Similarly, although significant improvements in anxiety and depression literacy were reported, Gulliver et al [37] found no significant effect of their Web-based interventions on help-seeking attitudes, intentions, or behaviors relative to controls (Multimedia Appendix 2).

Impact on Stigma

A multilingual Internet-based psychoeducational intervention was found to be effective in reducing personal but not perceived depression stigma [38] (Multimedia Appendix 2). Furthermore, reduction in depression stigma at postintervention and anxiety stigma at the 3-month follow-up was observed in the MHL and destigmatization condition of a brief, fully automated Internet-based help-seeking intervention [37] (Multimedia Appendix 2). Conversely, Taylor-Rodgers and Batterham [44] did not report significant changes in depression or suicide stigma, but a significant decrease in anxiety stigma (effect size=0.65) relative to the control group after a Web-based psychoeducational intervention (Multimedia Appendix 2). Likewise, Farrer et al [34] reported no overall significant improvement in depression stigma in response to MoodGYM/BluePages with or without participant follow-up; however, stigma appeared to be reduced after the intervention in both intervention conditions but only sustained in intervention without follow-up at 6 months. By the 12-month follow-up the

effect was not sustained in either intervention (Multimedia Appendix 2). Shandley et al [43] reported minor gender differences in stigmatizing attitudes, with women holding less stigmatizing attitudes compared with men, which were sustained from pre- to postintervention. Moreover, the intervention did not significantly change attitudes toward people with mental illness (Multimedia Appendix 2).

Impact on Mental Health Outcomes

Several studies reported improvements in mental health outcomes after Web-based interventions, in particular in studies that focused on patients with mental health problems. For example, Christensen et al [31] found both MoodGYM and BluePages to be effective in reducing depression symptomology in patients with clinical level of depression both in ITT analysis, which is more robust in accounting for missing data (d=0.4, 0.4, and 0.1 for MoodGYM, BluePages, and control, respectively), and in participants who completed all measures (n=414), as also reported by Griffiths et al [36]. Results were particularly encouraging in participants with CES-D scores of 16+ (n=369; completers: 0.6, 0.5, and 0.1 and completers with CES-D of 16+: 0.9, 0.75, and 0.25 for MoodGYM, BluePages, and control, respectively; Multimedia Appendix 2). Supporting these findings, Lintvedt et al [40] also found that MoodGYM and BluePages were effective in reducing depressive symptoms (d=0.57 on CES-D scores) and negative thoughts (d=0.50; Multimedia Appendix 2). Similarly, patients with schizophrenia experienced decrease in positive symptoms as a result of engaging with a targeted Web-based psychoeducation (d= −0.88) [41] (Multimedia Appendix 2).

Four studies reported no significant effect of their respective Web-based psychoeducational interventions on mental illness symptomology [32,38,44], nor was depressive symptomology significantly improved after a psychoeducational gaming intervention [43]. However, these studies were not specifically targeting individuals with mental health problems (Multimedia Appendix 2).

Discussion

Key Findings

The aims of this systematic review were twofold: first, to synthesize the existing evidence to establish the efficacy of Web-based interventions that seek to improve MHL. Second, to establish whether improvements in MHL translated into improvements in individual health seeking, reductions in stigmatizing attitudes toward people living with mental illness, and better health outcomes for individuals with mental illness. Outcomes from the review demonstrate that Web-based interventions targeting MHL are generally efficacious when they include the following "active" ingredients: the intervention comprises a structured program where participants are guided through a series of sequential steps, targets specific population or consumer groups, delivers evidence-based content (such as CBT and/or psychoeducation, depending on the target population), and is underpinned by a pedagogical approach that promotes interactivity and experiential learning. Examples of these types of interventions include MoodGYM and social network games [31,34,36,39,40]. Conversely, from our review

we observed that interventions that do not fully utilize the interactive potential of the Internet, and deliver generalist information to consumers using an unstructured, didactic approach, and/or where participants can navigate and access the website in any way they chose, are less successful in improving rates of MHL [32,43,44].

Several studies found positive associations between increased MHL and reduced symptomatology, especially for mild to moderate depression [31,36,40,41]. This suggests that Web-based interventions may be best suited to target people with less severe mental illness (eg, depression and anxiety) and that are of a mild to moderate nature (eg, mild to moderate depression rather than clinical depression). However, this finding should be cautiously interpreted as the studies making these findings included therapeutic components (such as CBT) alongside psychoeducational ones, and separating the effects of each is not possible.

Nevertheless, to date the most extensively tested interventions suited to people with mild to moderate depressive symptoms appear to be MoodGYM and BluePages. Initially developed and tested in Australia, the intervention has also been translated into Norwegian and tested in Norway [40]. These interventions have been rigorously evaluated using the "gold standard" RCT designs and generally reported improvements across a variety of measures of MHL [31,34,36,40] and symptomatology [31,36,40]. However, these interventions also have a high attrition rate on account of the time commitment required from participants (up to 6 hours), and researchers must carefully consider the merits of this approach in relation to their target population and particular mental illness. Interventions seeking to increase MHL in community members are unlikely to be successful using this intensive approach; likewise for patients whose mental illness precludes them from concentrating for long periods of time. As shown by Rotondi et al [41], there are other Internet-based interventions that may also hold promise for mental illnesses such as schizophrenia, yet this line of research requires further substantiation.

The relationship between increased MHL and reductions in stigmatizing attitudes is more complex. On the one hand, the evidence demonstrates a positive association between the two—as MHL increases, stigma decreases. On the other hand, this evidence is based on participants' self-report measures and it is difficult to establish how such attitudinal shifts inform everyday practices around inclusion and discrimination toward people with mental illness. Moreover, this review found no relationship between improvements in MHL and increased help seeking, suggesting that better knowledge about mental illness does not necessarily translate into people seeking the therapeutic care they might need. Avoiding the stigma of mental illness is one of the main reasons for not seeking appropriate and timely help [46,47]. Further research is needed to exemplify the potentially paradoxical relationship between MHL and help seeking.

Limitations of the Included Studies

Our findings are tempered by 4 limitations in the current evidence base. First, there was high variability between the studies on the duration of the exposure-response relationship.

Some studies incorporated a sustained engagement between the participants and the intervention into their design, and followed up over a prolonged period of time (eg, 12 months) to test the durability of the intervention (Tables 1 and 2). Other studies only had a one-off interaction between participants and the intervention and followed up participants for a very limited period (eg, 1 week; Tables 1 and 2). Second, monitoring participant adherence for complex interventions of this nature is challenging. Whether delivered via the web or through traditional platforms, there are many confounding factors—for example, social, cognitive, and structural—that could compromise the study results. Third, as several of the studies were complex interventions comprising multiple components it was unclear which components created the effects and whether these effects were intended or not. Finally, as acknowledged by many of the studies' authors, certain standard measures and techniques, such as ITT analysis, were not applied to the studies because of their small sample size. Thus the extent of the generalizability of many of the studies is not entirely clear.

Limitations of This Review

This review is also not without limitations. Only articles in English were included, thereby excluding research published in other languages. In addition, while the utmost care was taken to perform a thorough search, failure to include searches on specific mental illnesses (eg, schizophrenia) and literacy meant that we might have missed some studies, including evidence from the gray literature. Furthermore, because of the heterogeneity of the measures and outcomes of the included studies a meta-analysis could not be performed, limiting the overall rigor of the review. Lastly, as several of the studies were complex interventions comprising several components it proved difficult to disentangle which components influenced our target outcomes specifically.

Conclusions

To the best of our knowledge this is the first review to examine the efficacy of Web-based MHL interventions and to establish the relationship between these interventions and their effect on help seeking, stigma, and health outcomes. As our review demonstrates, there are several "active" ingredients to a successful Web-based intervention and, if properly implemented, these interventions can improve MHL and symptomatology among those with mild to moderate mental illness. Of greatest promise are the MoodGYM and BluePages interventions [31,34,36,40,41] that have proven to be not only efficacious but also cost-effective and culturally portable from Australia to Norway [40]. Future research could extend the utility of these interventions by testing their applicability in other country settings different from Australia.

Interestingly, much of the intervention research on MHL comes from Australia. For more than a decade now there have been several Australian public health campaigns addressing mental illness, nearly all of which have integrated components of MHL [10,48-52]. Australia is often cited as an exemplar in this field [13]. Much could also be learned from applying the lessons from Australian interventions in low- and middle-income countries, where Internet penetration is rapidly increasing [14]. Documenting how interventions are adopted and adapted to culturally diverse settings could open new horizons for scholarship vis-à-vis the relationship between MHL, help seeking, stigma, and health outcomes in culturally diverse settings. Finally, future studies could also explore how these relationships are influenced by the technology interface being used (eg, mobile phones vs computers). By realizing these future avenues for research, we can better harness the full potential of the Internet and new technologies in delivering new innovations to help improve the lives of people with mental illness.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Study overview and characteristics.

[PDF File (Adobe PDF File), 76KB - jmir_v18i6e165_app1.pdf]

Multimedia Appendix 2

Study outcomes.

[PDF File (Adobe PDF File), 77KB - jmir_v18i6e165_app2.pdf]

References

1. Steel Z, Marnane C, Iranpour C, Chey T, Jackson JW, Patel V, et al. The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013. Int J Epidemiol 2014 Apr;43(2):476-493. [doi: 10.1093/ije/dyu038] [Medline: 24648481]

2. Results from the 2011 National Survey on Drug Use and Health (NSDUH): Mental Health Findings, NSDUH Series H-45, HHS Publication No. (SMA) 12-4725. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2012. URL: http://media.samhsa.gov/data/NSDUH/2k11MH_FindingsandDetTables/2K11MHFR/NSDUHmhfr2011.htm [accessed 2015-12-17] [WebCite Cache ID 6iE2NCWuu]

3. Griffiths KM, Mendoza J, Carron-Arthur B. Whereto mental health reform in Australia: is anyone listening to our independent auditors? Med J Aust 2015 Mar 2;202(4):172-174. [Medline: 25716589]

4. Kohn R, Saxena S, Levav I, Saraceno B. The treatment gap in mental health care. Bull World Health Organ 2004 Nov;82(11):858-866 [FREE Full text] [Medline: 15640922]

5. Jorm AF. Mental health literacy: empowering the community to take action for better mental health. Am Psychol 2012 Apr;67(3):231-243. [doi: 10.1037/a0025957] [Medline: 22040221]

6. Kelly CM, Jorm AF, Wright A. Improving mental health literacy as a strategy to facilitate early intervention for mental disorders. Med J Aust 2007 Oct 1;187(7 Suppl):S26-S30. [Medline: 17908021]

7. Jorm AF, Korten AE, Jacomb PA, Christensen H, Rodgers B, Pollitt P. "Mental health literacy": a survey of the public's ability to recognise mental disorders and their beliefs about the effectiveness of treatment. Med J Aust 1997 Feb 17;166(4):182-186. [Medline: 9066546]

8. Jorm AF. Mental health literacy. Public knowledge and beliefs about mental disorders. Br J Psychiatry 2000 Nov;177:396-401 [FREE Full text] [Medline: 11059991]

9. Jorm AF, Christensen H, Griffiths KM. The impact of beyondblue: the national depression initiative on the Australian public's recognition of depression and beliefs about treatments. Aust N Z J Psychiatry 2005 Apr;39(4):248-254. [doi: 10.1111/j.1440-1614.2005.01561.x] [Medline: 15777361]

10. Yap MB, Reavley NJ, Jorm AF. Associations between awareness of beyondblue and mental health literacy in Australian youth: Results from a national survey. Aust N Z J Psychiatry 2012 Jun;46(6):541-552. [doi: 10.1177/0004867411435288] [Medline: 22679206]

11. Stuart H, Chen S, Christie R, Dobson K, Kirsh B, Knaak S, et al. Opening minds in Canada: targeting change. Can J Psychiatry 2014 Oct;59(10 Suppl 1):S13-S18 [FREE Full text] [Medline: 25565697]

12. Evans-Lacko S, Malcolm E, West K, Rose D, London J, Rüsch N, et al. Influence of Time to Change's social marketing interventions on stigma in England 2009-2011. Br J Psychiatry Suppl 2013 Apr;55:s77-s88 [FREE Full text] [doi: 10.1192/bjp.bp.113.126672] [Medline: 23553699]

13. Corrigan PW, Powell KJ, Al-Khouja MA. Examining the Impact of Public Service Announcements on Help Seeking and Stigma: Results of a Randomized Controlled Trial. J Nerv Ment Dis 2015 Nov;203(11):836-842. [doi: 10.1097/NMD.0000000000000376] [Medline: 26488913]

14. ICT Facts & Figures. Geneva, Switzerland: International Telecommunication Union; 2015. URL: https://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2015.pdf [accessed 2015-10-06] [WebCite Cache ID 6iE2i7VE3]

15. Gulliver A, Griffiths KM, Christensen H, Brewer JL. A systematic review of help-seeking interventions for depression, anxiety and general psychological distress. BMC Psychiatry 2012;12:81 [FREE Full text] [doi: 10.1186/1471-244X-12-81] [Medline: 22799879]

16. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg 2010;8(5):336-341 [FREE Full text] [doi: 10.1016/j.ijsu.2010.02.007] [Medline: 20171303]

17. National Institute for Health and Care Excellence. London, UK: National Institute for Health and Care Excellence; 2014. The Guidelines Manual URL: http://www.nice.org.uk/guidelinesmanual [accessed 2015-04-10] [WebCite Cache ID 6duulZQtJ]

18. Turk DC, Rudy TE, Sorkin BA. Neglected topics in chronic pain treatment outcome studies: determination of success. Pain 1993 Apr;53(1):3-16. [Medline: 8316386]

19. Brooks AC, Ryder D, Carise D, Kirby KC. Feasibility and effectiveness of computer-based therapy in community treatment. J Subst Abuse Treat 2010 Oct;39(3):227-235. [doi: 10.1016/j.jsat.2010.06.003] [Medline: 20667682]

20. Christensen H, Leach LS, Barney L, Mackinnon AJ, Griffiths KM. The effect of web based depression interventions on self reported help seeking: randomised controlled trial [ISRCTN77824516]. BMC Psychiatry 2006;6:13 [FREE Full text] [doi: 10.1186/1471-244X-6-13] [Medline: 16595018]

21. Collin PJ, Metcalf AT, Stephens-Reicher JC, Blanchard ME, Herrman HE, Rahilly K, et al. ReachOut.com: The role of an online service for promoting help-seeking in young people. Advances in Mental Health 2014 Dec 17;10(1):39-51. [doi: 10.5172/jamh.2011.10.1.39]

22. Hill W, Weinert C, Cudney S. Influence of a computer intervention on the psychological status of chronically ill rural women: preliminary results. Nurs Res 2006;55(1):34-42 [FREE Full text] [Medline: 16439927]

23. Hunkeler EM, Hargreaves WA, Fireman B, Terdiman J, Meresman JF, Porterfield Y, et al. A web-delivered care management and patient self-management program for recurrent depression: a randomized trial. Psychiatr Serv 2012 Nov;63(11):1063-1071. [doi: 10.1176/appi.ps.005332011] [Medline: 22983558]

24. Jorm AF, Kitchener BA, Fischer J, Cvetkovski S. Mental health first aid training by e-learning: a randomized controlled trial. Aust N Z J Psychiatry 2010 Dec;44(12):1072-1081. [doi: 10.3109/00048674.2010.516426] [Medline: 21070103]

25.   Jorm AF, Kitchener BA, Mugford SK. Experiences in applying skills learned in a Mental Health First Aid training course: a qualitative study of participants' stories. BMC Psychiatry 2005;5:43 [FREE Full text] [doi: 10.1186/1471-244X-5-43] [Medline: 16280088]

26.   Kobak KA, Stone WL, Wallace E, Warren Z, Swanson A, Robson K. A web-based tutorial for parents of young children with autism: results from a pilot study. Telemed J E Health 2011 Dec;17(10):804-808 [FREE Full text] [doi: 10.1089/tmj.2011.0060] [Medline: 22011005]

27.   Livingston JD, Tugwell A, Korf-Uzan K, Cianfrone M, Coniglio C. Evaluation of a campaign to improve awareness and attitudes of young people towards mental health issues. Soc Psychiatry Psychiatr Epidemiol 2013 Jun;48(6):965-973. [doi: 10.1007/s00127-012-0617-3] [Medline: 23124481]

28.   Meglic M, Furlan M, Kuzmanic M, Kozel D, Baraga D, Kuhar I, et al. Feasibility of an eHealth service to support collaborative depression care: results of a pilot study. J Med Internet Res 2010;12(5):e63 [FREE Full text] [doi: 10.2196/jmir.1510] [Medline: 21172765]

29.   Poole R, Simpson SA, Smith DJ. Internet-based psychoeducation for bipolar disorder: a qualitative analysis of feasibility, acceptability and impact. BMC Psychiatry 2012;12:139 [FREE Full text] [doi: 10.1186/1471-244X-12-139] [Medline: 22971042]

30.   Smith DJ, Griffiths E, Poole R, di FA, Barnes E, Kelly MJ, et al. Beating Bipolar: exploratory trial of a novel Internet-based psychoeducational treatment for bipolar disorder. Bipolar Disord 2011;13(5-6):571-577. [doi: 10.1111/j.1399-5618.2011.00949.x] [Medline: 22017225]

31.   Christensen H, Griffiths KM, Jorm AF. Delivering interventions for depression by using the internet: randomised controlled trial. BMJ 2004 Jan 31;328(7434):265 [FREE Full text] [doi: 10.1136/bmj.37945.566632.EE] [Medline: 14742346]

32.   Costin DL, Mackinnon AJ, Griffiths KM, Batterham PJ, Bennett AJ, Bennett K, et al. Health e-cards as a means of encouraging help seeking for depression among young adults: randomized controlled trial. J Med Internet Res 2009;11(4):e42 [FREE Full text] [doi: 10.2196/jmir.1294] [Medline: 19850549]

33.   Deitz DK, Cook RF, Billings DW, Hendrickson A. A web-based mental health program: reaching parents at work. J Pediatr Psychol 2009 Jun;34(5):488-494 [FREE Full text] [doi: 10.1093/jpepsy/jsn108] [Medline: 18845585]

34.   Farrer L, Christensen H, Griffiths KM, Mackinnon A. Web-based cognitive behavior therapy for depression with and without telephone tracking in a national helpline: secondary outcomes from a randomized controlled trial. J Med Internet Res 2012;14(3):e68 [FREE Full text] [doi: 10.2196/jmir.1859] [Medline: 22738715]

35.   Finkelstein J, Lapshin O. Reducing depression stigma using a web-based program. Int J Med Inform 2007 Oct;76(10):726-734. [doi: 10.1016/j.ijmedinf.2006.07.004] [Medline: 16996299]

36.   Griffiths KM, Christensen H, Jorm AF, Evans K, Groves C. Effect of web-based depression literacy and cognitive-behavioural therapy interventions on stigmatising attitudes to depression: randomised controlled trial. Br J Psychiatry 2004 Oct;185:342-349 [FREE Full text] [doi: 10.1192/bjp.185.4.342] [Medline: 15458995]

37.   Gulliver A, Griffiths KM, Christensen H, Mackinnon A, Calear AL, Parsons A, et al. Internet-based interventions to promote mental health help-seeking in elite athletes: an exploratory randomized controlled trial. J Med Internet Res 2012;14(3):e69 [FREE Full text] [doi: 10.2196/jmir.1864] [Medline: 22743352]

38.   Kiropoulos LA, Griffiths KM, Blashki G. Effects of a multilingual information website intervention on the levels of depression literacy and depression-related stigma in Greek-born and Italian-born immigrants living in Australia: a randomized controlled trial. J Med Internet Res 2011;13(2):e34 [FREE Full text] [doi: 10.2196/jmir.1527] [Medline: 21504872]

39.   Li TM, Chau M, Wong PW, Lai ES, Yip PS. Evaluation of a Web-based social network electronic game in enhancing mental health literacy for young people. J Med Internet Res 2013;15(5):e80 [FREE Full text] [doi: 10.2196/jmir.2316] [Medline: 23676714]

40.   Lintvedt OK, Griffiths KM, Sørensen K, Østvik AR, Wang CE, Eisemann M, et al. Evaluating the effectiveness and efficacy of unguided internet-based self-help intervention for the prevention of depression: a randomized controlled trial. Clin Psychol Psychother 2013;20(1):10-27. [doi: 10.1002/cpp.770] [Medline: 21887811]

41.   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]

42.   Roy MJ, Taylor P, Runge W, Grigsby E, Woolley M, Torgeson T. Web-based post-traumatic stress disorder education for military family members. Mil Med 2012 Mar;177(3):284-290. [Medline: 22479915]

43.   Shandley K, Austin D, Klein B, Kyrios M. An evaluation of 'Reach Out Central': an online gaming program for supporting the mental health of young people. Health Educ Res 2010 Aug;25(4):563-574 [FREE Full text] [doi: 10.1093/her/cyq002] [Medline: 20150530]

44.   Taylor-Rodgers E, Batterham PJ. Evaluation of an online psychoeducation intervention to promote mental health help seeking attitudes and intentions among young adults: randomised controlled trial. J Affect Disord 2014 Oct 15;168:65-71. [doi: 10.1016/j.jad.2014.06.047] [Medline: 25038293]

45.   Gupta SK. Intention-to-treat concept: A review. Perspectives in Clinical Research 2011;2(3):109-112.

46.   Corrigan PW, Druss BG, Perlick DA. The Impact of Mental Illness Stigma on Seeking and Participating in Mental Health Care. Psychol Sci Public Interest 2014 Oct;15(2):37-70. [doi: 10.1177/1529100614531398] [Medline: 26171956]

XSL•FO
RenderX

47. Corrigan P. Association for Psychological Sciences. Washington DC: Association for Psychological Sciences; 2014. Stigma as a Barrier to Mental Health Care URL: http://www.psychologicalscience.org/index.php/news/releases/stigma-as-a-barrier-to-mental-health-care.html [accessed 2015-09-01] [WebCite Cache ID 6duuH4N2p]

48. Highet NJ, Luscombe GM, Davenport TA, Burns JM, Hickie IB. Positive relationships between public awareness activity and recognition of the impacts of depression in Australia. Aust N Z J Psychiatry 2006 Jan;40(1):55-58. [doi: 10.1111/j.1440-1614.2006.01742.x] [Medline: 16403039]

49. Jorm AF. Australian young people's awareness of headspace, beyondblue and other mental health organizations. Australas Psychiatry 2009 Dec;17(6):472-474. [doi: 10.1080/10398560902915596] [Medline: 20001369]

50. Jorm AF, Christensen H, Griffiths KM. The impact of beyondblue: the national depression initiative on the Australian public's recognition of depression and beliefs about treatments. Aust N Z J Psychiatry 2005 Apr;39(4):248-254. [doi: 10.1111/j.1440-1614.2005.01561.x] [Medline: 15777361]

51. Jorm AF, Christensen H, Griffiths KM. Changes in depression awareness and attitudes in Australia: the impact of beyondblue: the national depression initiative. Aust N Z J Psychiatry 2006 Jan;40(1):42-46. [doi: 10.1111/j.1440-1614.2006.01739.x] [Medline: 16403036]

52. Jorm A. Mental health literacy: empowering the community to take action for better mental health. Am Psychol 2012 Apr;67(3):231-243. [doi: 10.1037/a0025957] [Medline: 22040221]

Abbreviations

CBT: cognitive behavioral therapy
CES-D: Center for Epidemiologic Studies-Depression
ITT: intention-to-treat
MHL: mental health literacy
NICE: National Institute for Health and Care Excellence
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PTSD: posttraumatic stress disorder
RCT: randomized controlled trial

Original Paper

Trajectories of Suicidal Ideation in People Seeking Web-Based Help for Suicidality: Secondary Analysis of a Dutch Randomized Controlled Trial

Trine Madsen[1], PhD; Bregje van Spijker[2], PhD; Karen-Inge Karstoft[3], PhD; Merete Nordentoft[1], PhD, Prof. Dr; Ad JFM Kerkhof[4], PhD, Prof.

[1]Copenhagen Mental Health Center, Copenhagen University Hospital, Hellerup, Denmark

[2]National Institute for Mental Health Research, Research School of Population Health, The Australian National University, Canberra, Australia

[3]Research and Knowledge Centre, The Danish Veteran Centre, Ringsted, Denmark

[4]Department of Clinical, Neuro, and Developmental Psychology and the EMGO, Institute for Health and Care Research, Faculty of Behavioural and Movement Sciences, Vrije Universiteit, Amsterdam, Netherlands

Corresponding Author:
Trine Madsen, PhD
Copenhagen Mental Health Center
Copenhagen University Hospital
Kildegårdsvej 28, opgang 15, 4.sal
Hellerup, 2900
Denmark
Phone: 45 35316236
Fax: 45 38647459
Email: trine.madsen@regionh.dk

Abstract

Background: Suicidal ideation (SI) is a common mental health problem. Variability in intensity of SI over time has been linked to suicidal behavior, yet little is known about the temporal course of SI.

Objective: The primary aim was to identify prototypical trajectories of SI in the general population and, secondarily, to examine whether receiving Web-based self-help for SI, psychiatric symptoms, or sociodemographics predicted membership in the identified SI trajectories.

Methods: We enrolled 236 people, from the general Dutch population seeking Web-based help for SI, in a randomized controlled trial comparing a Web-based self-help for SI group with a control group. We assessed participants at inclusion and at 2, 4, and 6 weeks. The Beck Scale for Suicide Ideation was applied at all assessments and was included in latent growth mixture modeling analysis to empirically identify trajectories.

Results: We identified 4 SI trajectories. The high stable trajectory represented 51.7% (122/236) of participants and was characterized by constant high level of SI. The high decreasing trajectory (50/236, 21.2%) consisted of people with a high baseline SI score followed by a gradual decrease to a very low score. The third trajectory, high increasing (12/236, 5.1%), also had high initial SI score, followed by an increase to the highest level of SI at 6 weeks. The fourth trajectory, low stable (52/236, 22.0%) had a constant low level of SI. Previous attempted suicide and having received Web-based self-help for SI predicted membership in the high decreasing trajectory.

Conclusions: Many adults experience high persisting levels of SI, though results encouragingly indicate that receiving Web-based self-help for SI increased membership in a decreasing trajectory of SI.

KEYWORDS

suicidal ideation; online self-help; trajectories; latent growth mixture modeling

XSL•FO
RenderX

Introduction

Suicidal ideation (SI) is a common mental health problem. The lifetime prevalence of SI in Western countries has been reported to be around 10% [1,2]. For a person with SI, the prevalence of having attempted suicide is as high as 29% [1]. Variability in intensity of SI over time has previously been linked to suicidal behavior [3,4], yet remarkably little is known about the temporal course in adults with SI.

Traditionally, longitudinal studies have reported the prevalence of self-reported SI at several time points or provided the mean change in SI over time for a whole study cohort or in predefined categories based on, for instance, diagnosis [5-8]. While this has provided valuable knowledge, these studies have not examined trends in individual variability of SI. To our knowledge, only a few studies have examined prototypical trajectories of SI based on individual variation in frequency and course of SI over time. These studies have been carried out in certain predefined subsamples, such as psychiatric patient samples [4,9-12] or in teenage samples [13-16]. Hence, little is known about the individual variation of SI in the adult general population over time.

By applying data from a longitudinal Dutch randomized controlled trial comparing a group following a Web-based self-help program for suicidal thoughts with a waitlist control group (Netherlands Trial Register NTR1689), we had the opportunity to examine prototypical patterns of SI in an adult sample from the general population seeking Web-based help for SI. Promising results from this randomized controlled trial have previously been published showing that SI was significantly more reduced in the Web-based self-help program receivers than in the control participants [7]. Yet, while the traditionally applied statistics in that study showed an overall small significant mean effect size of the program (see Figure 1), that study did not evaluate how many trial participants experienced decreasing, increasing, stable, or fluctuating SI over the study period. A better understanding of this individual variation in temporal patterns of SI and what characterizes these patterns may help target preventive actions toward those at greatest risk of suicidal behavior.

Our primary aim was to identify prototypical trajectories of SI, thereby increasing current knowledge on individual variability of SI. Secondarily, we examined whether baseline sociodemographics, clinical symptoms, or the Web-based self-help program was associated with membership in the identified SI trajectories.

Figure 1. Mean change in suicidal ideation as assessed by the Beck Scale for Suicide Ideation (BSS), overall and by randomization group.
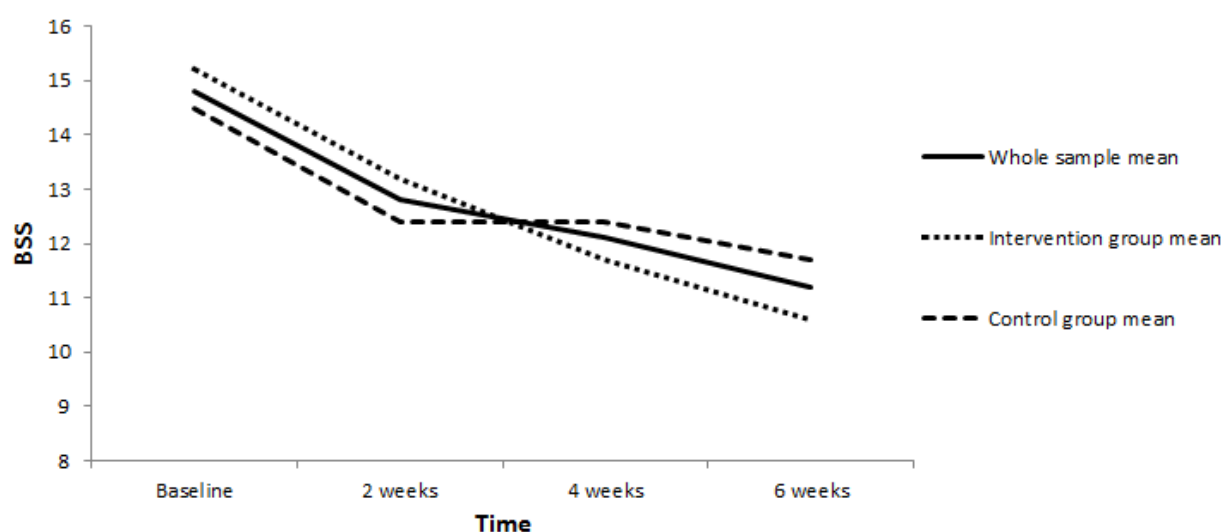
Methods

Full details of the methods of the randomized controlled trial have previously been described [7,17]. Here, we present details of relevance to this study.

Procedure

We recruited participants from the general population through newspaper advertisements and banners on the Internet that directed interested people to a website where they obtained information about the study and were able to register. Inclusion in the study required a minimum age of 18 years, access to the Internet and an email address, being fluent in Dutch, having

mild to moderate SI (defined as a score between 1 and 26 on the Beck Scale for Suicide Ideation, BSS) [18], and not being severely depressed (defined as a score >39 on the Beck Depression Inventory, BDI) [19]. We determined these cutoff scores in consultation with clinical experts. An independent researcher using a block design randomly allocated participants to either the intervention or the control condition.

Because this study was conducted in a vulnerable population, we used safety procedures [7,17]. Each time a participant exceeded a cutoff score on SI (BSS >26) or depressive symptoms (BDI >39), we carried out a risk assessment over the phone. If deemed necessary, or if a participant could not be reached, we contacted their general practitioner. The study was

approved by the Medical Ethics Committee of the VU University Medical Centre (registration number 2008/204).

Intervention

The Web-based self-help program was based on elements from cognitive behavioral therapy, dialectical behavioral therapy, problem-solving therapy, and mindfulness-based cognitive therapy, which have all been shown to reduce suicidality [20-23]. The program consisted of 6 modules for participants to work through independently, ideally doing 1 module per week and spending approximately 30 minutes daily on the module.

Each module contained theory and core exercises. Module 1 aimed at helping participants recognize how often they repeat suicidal thoughts and learning to manage this worrying or ruminating repetition better. Module 2 focused on learning to tolerate and regulate intense emotions. The theory section explained how to recognize an upcoming crisis and tapped into dealing with the urge to self-harm. Core exercises introduced different ways of coping with intense emotions, such as behavioral activation (eg, seeking distraction) and acceptance (waiting until the feelings subside). Participants were also encouraged to make a crisis plan. Modules 3 to 5 dealt with identifying automatic thoughts, recognizing common thinking patterns (all-or-nothing thinking, overgeneralization, or mind reading), and cognitively restructuring the three most important identified negative automatic thoughts. Module 6 was dedicated to preventing relapse and discussed the possibility of future setbacks and disappointments.

The control group received access to a website with information on suicidality, that is, how common it is and its risk factors, and provided a list of common places to seek treatment for suicidality. The control group was provided with access to the self-help program at the 6-week follow-up.

Measures

Our primary measure, which we used to identify latent trajectories of SI, was the BSS questionnaire, which we distributed to participants at baseline, at weeks 2 and 4 of the intervention, and finally at the sixth and last week of the intervention. Furthermore, we administered questions on the following at baseline, and subsequently applied and examined them as possible predictors of trajectory membership: sex, age, living with partner (yes/no), paid employment (yes/no), having children (yes/no), random allocation group (control/Web-based self-help intervention), and currently receiving other help such as psychiatric or psychological therapy (yes/no). We also examined clinical factors such as having attempted suicide before baseline (item from the BSS), depression symptoms measured by the BDI, and levels of hopelessness as measured by the Beck Hopelessness Scale [24].

Statistical Methods

We applied latent growth mixture modelling (LGMM) to estimate trajectories of SI. This data-driven statistical method identifies subgroups in a sample based on shared growth parameters (ie, intercept and slope). As such, in LGMM it is assumed that multiple subpopulations exist in a sample; however, no assumptions are made about the number of subpopulations or their specific growth parameters. Hence, individuals are classified into possible unobserved subgroups based on common profile patterns and, subsequently, between-group differences can be examined [25,26]. We included the level of self-reported SI collected at all 4 assessments (baseline, and 2, 4, and 6 weeks) and we used this information to model unique latent trajectories of SI. In general, attrition at follow-up was low: 6.8% (220/236, 2-week assessment), 10.6% (211/236, 4-week assessment), and 8.9% (215/236, 6-week assessment). We included all patients (N=236) who answered questions on SI at baseline in the LGMM and handled missing data by applying the full information maximum likelihood method [27]. We estimated a series of LGMM models with the number of classes ranging from 1 to 6. We used the fit estimates of Bayesian information criteria (BIC), Akaike information criteria (AIC), and sample size-adjusted BIC (adj.BIC) to evaluate the models. Specifically, lower values of these indexes imply a better model fit. Furthermore, we tested improvement of model fit by adding an extra class by estimating the Lo-Mendell-Rubin adjusted likelihood ratio test, the Vuong-Lo-Mendell-Rubin likelihood ratio test, and the bootstrap likelihood ratio test. We also estimated entropy, which is an estimate of classification accuracy and describes the probability of the model assigning individuals correctly in trajectories. According to the traditional approach, however, model selection is based on a combination of factors in addition to fit indices, namely parsimony of the model, theoretical justification, and interpretability [25]. In the first step, we estimated trajectories without including covariates. Second, we tested all covariates univariately for association with the most likely class membership using Pearson chi-square test (categorical variable) or analysis of variance F test (continuous variables). Third, we included all covariates with a univariate $P<.2$ in the 3-step approach suggested by Asparouhov and Muthén [28] to conduct multivariable testing of predictors of trajectory membership. When using this approach, covariates are not initially included in the LGMM but treated as auxiliary variables, that is, covariates do not influence the formation of classes. However, the probabilistic nature of class membership assignment is still accounted for. Thus, we first established latent classes, and then examined covariates as possible predictors of membership in identified latent classes in multivariable-adjusted analyses [28]. The analyses were carried out in Mplus version 7 (Muthén & Muthén) and in IBM SPSS version 22 (IBM Analytics).

Results

Table 1 presents goodness of fit statistics for the LGMM analyses. We estimated models from 1 to 6 classes with fixed variance around the quadratic term and the slope, but with a freely varying intercept.

Table 1. Goodness of fit statistics for 1- to 6-class solutions and *P* values for test of class models.

No. of classes	Fit estimates[a]			*P* values[b]			Entropy[c]
	AIC[d]	BIC[e]	adj.BIC[f]	Vuong-Lo- Mendell-Rubin likelihood ratio test	Lo-Mendell- Rubin adjusted likelihood ratio test	Bootstrap likelihood ratio test	
1	5798	5825	5800				
2	5742	5783	5745	<.01	<.01	<.01	.69
3	5709	5764	5714	.23	.24	<.01	.77
4	5673	5742	5679	.06	.07	<.01	.85
5	5647	5730	5654	.24	.26	<.01	.86
6	5630	5727	5638	.08	.10	<.01	.87

[a]Lower values of AIC, BIC, and adj.BIC indicate better model fit.

[b]*P* values for test against model minus 1 class.

[c]Entropy estimates ranges from 0 to 1 and assess the accuracy with which models classify individuals into their most likely class; higher scores represent greater classification accuracy.

[d]Akaike information criteria.

[e]Bayesian information criteria.

[f]Sample size-adjusted BIC.

We obtained the most parsimonious model with the 4-class solution, where fit estimates (AIC, BIC, and adj.BIC) were low, the bootstrap likelihood ratio test performed with a significant *P*<0000, and both the Vuong-Lo-Mendell-Rubin likelihood ratio test and the Lo-Mendell-Rubin adjusted likelihood ratio test indicated borderline *P*=.06. Fit estimates did not decrease substantially with the addition of extra classes after the 4-class solution and, in addition, the sample size in one of the classes in the 5-class solution was very low (n=4). Further, the 5-class model performed with an insignificant Vuong-Lo-Mendell-Rubin likelihood ratio test, and the Lo-Mendell-Rubin likelihood ratio tests had *P*>.2. Based on these results, we chose the 4-class over the 5-class model. In addition, the 4-class solution had high entropy (.85) and was clinically plausible, so we settled on this model as the optimal representation of SI in our data.

Of the 4 identified trajectories, 3 had high intercepts (BSS>14), indicating that individuals in these latent classes had high intensity or frequency of SI at baseline (see Figure 2). These 3 classes then developed into different patterns. The largest class, named high stable, represented 51.7% (122/236) of the participants and was characterized by a constant high score on the BSS throughout the 6-week study period. The second class, high decreasing, included 21.2% (50/236) of the sample and consisted of people with a high baseline SI score (BSS=16), which gradually decreased to an average BSS score of 2.1 at the final follow-up at week 6. Participants in the third class, high increasing, also had a high BSS score of 14 at baseline, followed by an increase throughout the study period to the highest level of SI at 6 weeks. This class consisted of 5.1% (12/236) of the participants. The fourth, low stable, consisted of 22.0% (52/236) of the participants and differed from the other

3 classes by having a considerably lower average level of SI at baseline (BSS=4.7). This level of SI was constantly low through the final assessment.

Table 2 presents characteristics by trajectory membership and univariate *P* tests of differences between them. Random allocation group, partner status, a history of attempted suicide, and higher hopelessness and depression scores indicated significant between-trajectory differences. Along with status of employment, which was nearly significant, we introduced all these variables into the multivariate predictor analyses of trajectory membership presented in Table 3 and Table 4. Participants who had received the intervention (odds ratio, OR 3.15, 95% CI 1.19–7.67) or lived with a partner (OR 3.12, 95% CI 1.27–7.67) had a 3-fold higher likelihood of membership in the high decreasing class than of membership in the high stable trajectory (see Table 3). Furthermore, relative to the high stable class, a history of attempted suicide predicted membership in the high decreasing class (OR 2.72, 95% CI 1.10–6.70). Similarly, members of the high stable class differed from members of the low stable class, who had a significantly lower level of hopelessness (OR 0.78, 95% CI 0.67–0.91). When we used the low stable class as the reference group (see Table 4), we found that participants receiving Web-based self-help (OR 3.84, 95% CI 1.22–12.12), those with a history of attempted suicide (OR 3.68, 95% CI 1.05–13.0), and those with higher levels of hopelessness (OR 1.24, 95% CI 1.01–1.52) were significantly more likely to be members of the high decreasing class. We found no significant differences in characteristics between members of the high increasing class and the high decreasing class, and therefore Table 3 or Table 4 do not display predictors between those 2 classes.

Table 2. Baseline characteristics of members in each suicidal ideation trajectory.

Characteristics	Latent trajectory[a]				P value
	High decreasing (n=46)	High increasing (n=11)	High stable (n=128)	Low stable (n=51)	
Age in years, mean (SD)	41 (12.3)	36 (12.0)	41 (14.3)	41 (13.8)	.66[b]
Sex, n (%)					.96[c]
Women	30 (65)	8 (73)	85 (66)	33 (65)	
Men	16 (35)	3 (27)	43 (34)	18 (35)	
Having children, n (%)					.58[c]
No	27 (60)	9 (82)	78 (61)	31 (63)	
Yes	18 (40)	2 (18)	49 (39)	18 (37)	
Random allocation, n (%)					.04[c]
Control group	15 (33)	6 (55)	68 (53)	31 (61)	
Intervention group	31 (67)	5 (45)	60 (47)	20 (39)	
Receiving other help, n (%)					.25[c]
No	21 (47)	2 (18)	52 (42)	25 (50)	
Yes	24 (53)	9 (82)	73 (58)	25 (50)	
Paid employment, n (%)					.07[c]
No	19 (42)	9 (82)	67 (53)	21 (43)	
Yes	26 (58)	2 (18)	60 (47)	28 (57)	
Living with partner, n (%)					.03[c]
No	20 (44)	9 (82)	88 (65)	29 (57)	
Yes	26 (57)	2 (18)	45 (35)	22 (43)	
History of attempted suicide, n (%)					<.01[c]
No	21 (47)	2 (18)	78 (61)	36 (74)	
Yes	24 (53)	9 (82)	49 (39)	13 (27)	
Clinical scale, mean score (SD)					
Beck Depression Inventory	26 (9.9)	34 (12.0)	29 (8.3)	22 (8.6)	<.01[b]
Beck Hopelessness Scale	15 (3.4)	15 (3.4)	15 (3.5)	12 (3.6)	<.01[b]

[a]Numbers of members in each class diverge slightly from those in Figure 2 because, while the figure is based on posterior probabilities, classes used for post hoc analyses are based on most likely class membership.

[b]Analysis of variance *F* test.

[c]Pearson chi-square test.

Table 3. Odds ratios (95% CI) from multivariable logistic regression analyses describing the association between baseline characteristics and membership in suicidal ideation trajectories using the high stable trajectory as the reference.

Characteristics	Low stable	High increasing	High decreasing
Random allocation			
Control group	1	1	1
Intervention group	0.82 (0.31–2.13)	0.66 (0.11–3.78)	3.15 (1.19–7.67)*
Paid employment			
No	1	1	1
Yes	1.19 (0.48–2.97)	0.18 (0.01–4.25)	1.30 (0.52–3.24)
Living with partner			
No	1	1	1
Yes	1.79 (0.68–4.70)	0.57 (0.08–4.08)	3.12 (1.27–7.67)*
History of attempted suicide			
No	1	1	1
Yes	0.74 (0.24–2.26)	8.64 (0.16–452.7)	2.72 (1.10–6.70)*
Clinical scale			
Beck Depression Inventory	0.95 (0.89–1.01)	1.06 (0.95–1.19)	0.95 (0.90–1.01)
Beck Hopelessness Scale	0.78 (0.67–0.91) **	0.87 (0.68–1.11)	0.96 (0.81–1.15)

* $P<.05$.
** $P<.01$.

Table 4. Odds ratios (95% C) from multivariable logistic regression analyses describing the association between baseline characteristics and membership in suicidal ideation trajectories using the low stable trajectory as the reference.

Characteristics	High increasing	High decreasing
Random allocation		
Control group	1	1
Intervention group	0.8 (0.12–5.25)	3.84 (1.22–12.12)*
Paid employment		
No	1	1
Yes	0.15 (0.01–3.48)	1.09 (0.36–3.31)
Living with partner		
No	1	1
Yes	0.32 (0.04–2.36)	1.74 (0.58–5.21)
History of attempted suicide		
No	1	1
Yes	11.7 (0.26–516.8)	3.68 (1.05–13.0)*
Clinical scale		
Beck Depression Inventory	1.12 (0.99–1.26)	1.0 (0.93–1.08)
Beck Hopelessness Scale	1.11 (0.86–1.45)	1.24 (1.01–1.52)*

* $P<.05$.

Figure 2. Trajectories of suicidal ideation as assessed by Beck Scale for Suicide Ideation (BSS) scores.
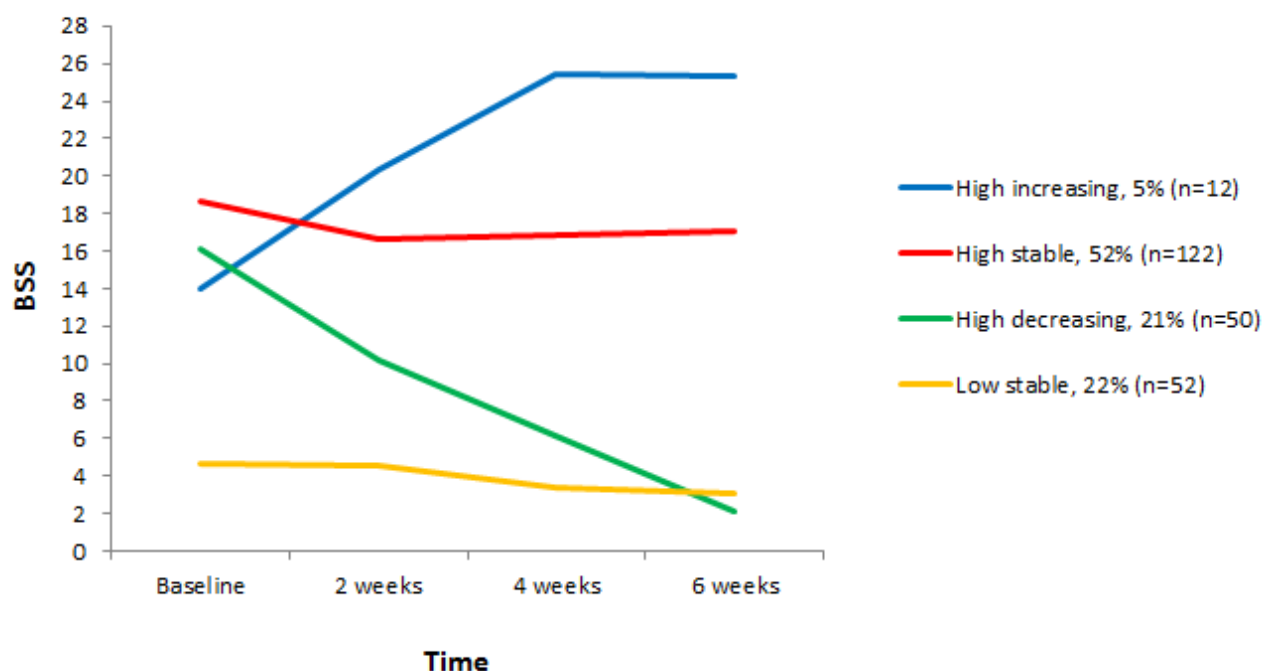
Discussion

This study aimed to examine prototypical trajectories of SI in an adult population seeking Web-based treatment for suicidal thoughts. The results of the LGMM confirmed heterogeneity by identifying 4 distinct trajectories of SI in this population. A notable finding was that those experiencing a decrease in SI (members of the high decreasing class) differed in characteristics from those with a more stable level of SI (both high stable and low stable) during the study period. Those receiving the Web-based intervention (vs the control group) and having a history of attempted suicide were more than 3 times as likely to be members of the high decreasing class as to be members of both stable classes. Members of the high decreasing class were also more likely than members of the high stable class to be living with a partner, and they had a higher level of hopelessness at baseline than did members of the low stable class. Overall, the findings of this study yield important and new information about developmental patterns of SI and especially how these patterns associate with the Web-based self-help intervention.

Trajectories of SI

This study has demonstrated that patterns of SI vary greatly, which is not detected using traditional analyses. While traditional analyses showed an overall small significant mean effect size in SI (Figure 1), the LGMM revealed that this effect was minimal for 74% of participants (low stable plus high stable classes), large for 21% of participants (high decreasing) who experienced a great decrease in SI, and reversed for 5% of participants who developed higher levels of SI. Because SI is an important risk factor for suicidal behaviour, it is encouraging to find that SI decreased from a worrisomely high, from a

clinician's standpoint, baseline level (BSS=16) to a considerably lower level (BSS=2) in one-fifth of participants, and that being in the intervention group was a significant predictor for membership in this class. Less reassuring is that 74% (high stable plus low stable classes) kept a stable level of SI, though it can be argued that the 22% in the low stable class had only limited potential to decrease their levels of SI. Still, the high stable class represents the majority of participants, indicating that SI persists in many despite over half in this class reporting receiving care from a professional at baseline. This indicates that SI persists (at least over a 6-week period) in many in spite of getting professional help. Although studies examining heterogeneity generally identify at least one trajectory of stable high SI, this class usually includes between 6% and 20% of participants, with one study reporting as many as 40% of participants falling into that class [9-16]. Our uncommonly large percentage of the sample with high ongoing SI may be explained by the short 6-week period compared with periods of a year or more in most other studies [10,11,13-16]. Another reason for this could be that our sample consisted of people who were required to have some level of SI at baseline, whereas the other studies examined heterogeneity in SI in samples of teenagers or psychiatric patients whether or not they reported SI [11-16]. Furthermore, a variety of instruments to measure SI was used across studies examining trajectories in SI, making comparison more challenging. Finally, the high percentage of participants in our high stable class may be illustrative of a group of chronically suicidal people who are desperate for help but are not getting what they need in the current health care system (nor in the current self-help intervention). More research will be needed to study this group and their needs.

Characteristics of SI Trajectory Class Members

We also identified predictors of membership in trajectories. Most important, membership in the high decreasing trajectory, compared with the 2 stable trajectories, was significantly associated with having received the Web-based intervention program, indicating that the program had a reducing effect on SI. This is consistent with earlier publications on this sample, which showed a small but significant overall mean effect size (a reduction of 4.6 points on the BSS) [7,29], but adds to this knowledge by showing that the Web-based program seems to have a larger effect (a reduction of 14 points on the BSS), but in a smaller subsample. Other factors that predicted membership in this class included previous suicidal behavior, living with a partner, and hopelessness. While this is speculative, people with a history of suicidal behavior may be more motivated to change their life situation and consequently would benefit more from a Web-based self-help program aimed at reducing SI. Furthermore, regarding motivation, living with a partner was also significantly more likely in the high decreasing class than in the high stable class, and having a partner may provide emotional support and help motivate their partner with SI to improve in mental well-being and health. Finally, having feelings of hopelessness is a well-known risk factor for suicidality, and the fact that members of the high decreasing class scored significantly lower on baseline hopelessness than those in the high stable trajectory is consistent with the existing literature on this [30].

Strength and Weaknesses

The high follow-up response rate strengthens this study, as this makes the estimation of SI trajectories more precise. Furthermore, we measured SI with a validated scale (BSS), as opposed to a single-item question on SI, which was applied in some of the other studies of heterogeneity in SI [9,11,15]. An obvious methodological concern is the low power of this study, with only 236 participants. This is especially a problematic issue when 1 trajectory (the high increasing) had very few members, as the power to predict membership in that class is low. Had the sample been larger, we would have had the power to detect predictors of membership in the high increasing class, which obviously is an important group of people to characterize and monitor, as their level of SI increased during the study period. Table 2 shows that most participants in the high increasing trajectory were already receiving other help, had attempted suicide before, had higher BDI scores, and were single or unemployed. Hence, this group of participants was characterized by known risk factors of suicidality. Based on this and because less than half of the members of this trajectory had received the intervention, we have reasons to believe that the worsening course of SI in this group probably does not reflect an iatrogenic effect of the Web-based self-help.

Conclusion

The analyses of trajectories of SI have provided clinically interesting information. In particular, it was notable that being assigned to the Web-based self-help program, having a history of suicidal behavior, having a partner, and having lower baseline levels of hopelessness were associated with experiencing a large decrease in SI during the 6-week study period, even though this applied to a small proportion of the sample. A discouraging finding was that a large proportion of the Web-based help seekers experienced persistently high levels of SI, and future research should address whether this group might benefit more from tailored Web-based programs or from face-to-face help. Recently, this Web-based self-help program to reduce SI was found to be cost effective too [29], so it is a promising bid for an intervention that can be offered to many via Web-based technology, although future studies are still needed to confirm a long-term effect of the program on SI.

Conflicts of Interest

The authors report no financial or other relationship relevant to the subject of this paper. BvS and AK are authors of the Web-based self-help program described in this paper. BvS and AK receive royalties from an adapted paper version of the self-help program described in this paper under the title "Piekeren over zelfdoding" (in Dutch), published by Boom, Amsterdam (2012).

References

1. Nock M, Borges G, Bromet E, Alonso J, Angermeyer M, Beautrais A, et al. Williams Dl. Cross-national prevalence and risk factors for suicidal ideation, plans and attempts. Br J Psychiatry 2008 Feb;? 2008 Feb;192(2):98-105. [doi: 10.1192/bjp.bp.107.040113]
2. ten Have M, de Graaf R, van Dorsselaer S, Verdurmen J, van't Land H, Vollebergh W. Suïcidaliteit in de Algemene Bevolking: Gedachten en Pogingen. Utrecht, the Netherlands: Ladenius Communicate BV; 2006.

XSL•FO

RenderX

3. Witte T, Fitzpatrick K, Joiner T, Schmidt N. Variability in suicidal ideation: a better predictor of suicide attempts than intensity or duration of ideation? J Affect Disord 2005 Oct;88(2):131-136. [doi: 10.1016/j.jad.2005.05.019] [Medline: 16054227]

4. Prinstein M, Nock M, Simon V, Aikins J, Cheah C, Spirito A. Longitudinal trajectories and predictors of adolescent suicidal ideation and attempts following inpatient hospitalization. J Consult Clin Psychol 2008;76(1):92-103. [doi: 10.1037/0022-006X.76.1.92]

5. Selby EA, Yen S. Six-month trajectory of suicidal ideation in adolescents with borderline personality disorder. Suicide Life Threat Behav 2014 Feb;44(1):89-100 [FREE Full text] [doi: 10.1111/sltb.12057] [Medline: 24112120]

6. King CA, Horwitz A, Berona J, Jiang Q. Acutely suicidal adolescents who engage in bullying behavior: 1-year trajectories. J Adolesc Health 2013 Jul;53(1 Suppl):S43-S50 [FREE Full text] [doi: 10.1016/j.jadohealth.2012.09.016] [Medline: 23790201]

7. van Spijker BAJ, van Straten A, Kerkhof AJFM. Effectiveness of online self-help for suicidal thoughts: results of a randomised controlled trial. PLoS One 2014;9(2):e90118 [FREE Full text] [doi: 10.1371/journal.pone.0090118] [Medline: 24587233]

8. Lee C, Seo D. Trajectory of suicidal ideation in relation to perceived overweight from adolescence to young adulthood in a representative United States sample. J Adolesc Health 2013 Dec;53(6):712-716. [doi: 10.1016/j.jadohealth.2013.06.013] [Medline: 23910569]

9. Restifo E, Kashyap S, Hooke G, Page A. Daily monitoring of temporal trajectories of suicidal ideation predict self-injury: A novel application of patient progress monitoring. Psychother Res 2015;25(6):705-713. [doi: 10.1080/10503307.2015.1006707]

10. Czyz E, King CA. Longitudinal trajectories of suicidal ideation and subsequent suicide attempts among adolescent inpatients. J Clin Child Adolesc Psychol 2015;44(1):181-193 [FREE Full text] [doi: 10.1080/15374416.2013.836454] [Medline: 24079705]

11. Madsen T, Karstoft K, Secher RG, Austin S, Nordentoft M. Trajectories of suicidal ideation in patients with first-episode psychosis: secondary analysis of data from the OPUS trial. Lancet Psychiatry 2016 May;3(5):443-450. [doi: 10.1016/S2215-0366(15)00518-0] [Medline: 26948484]

12. Kasckow J, Youk A, Anderson SJ, Dew MA, Butters MA, Marron MM, et al. Trajectories of suicidal ideation in depressed older adults undergoing antidepressant treatment. J Psychiatr Res 2016 Feb;73:96-101. [doi: 10.1016/j.jpsychires.2015.11.004] [Medline: 26708830]

13. Musci RJ, Hart SR, Ballard ED, Newcomer A, Van Eck K, Ialongo N, et al. Trajectories of suicidal ideation from xixth through tenth grades in predicting suicide attempts in young adulthood in an urban African American cohort. Suicide Life Threat Behav 2016 Jun;46(3):255-265. [doi: 10.1111/sltb.12191] [Medline: 26395337]

14. Adrian M, Miller A, McCauley E, Vander SA. Suicidal ideation in early to middle adolescence: sex-specific trajectories and predictors. J Child Psychol Psychiatry 2016 May;57(5):645-653. [doi: 10.1111/jcpp.12484] [Medline: 26610726]

15. Nkansah-Amankra S. Adolescent suicidal trajectories through young adulthood: prospective assessment of religiosity and psychosocial factors among a population-based sample in the United States. Suicide Life Threat Behav 2013 Aug;43(4):439-459 [FREE Full text] [doi: 10.1111/sltb.12029] [Medline: 23601148]

16. Giletta M, Prinstein MJ, Abela JRZ, Gibb BE, Barrocas AL, Hankin BL. Trajectories of suicide ideation and nonsuicidal self-injury among adolescents in mainland China: Peer predictors, joint development, and risk for suicide attempts. J Consult Clin Psychol 2015 Apr;83(2):265-279. [doi: 10.1037/a0038652] [Medline: 25688812]

17. van Spijker BA, van Straten A, Kerkhof AJ. The effectiveness of a web-based self-help intervention to reduce suicidal thoughts: a randomized controlled trial. Trials 2010;11:25 [FREE Full text] [doi: 10.1186/1745-6215-11-25] [Medline: 20214777]

18. Beck A, Steer R. Manual for the Beck Scale for Suicide Ideation. San Antonio, TX: The Psychological Corporation; 1991.

19. Beck A, Steer R. Beck Depression Inventory: Second Edition Manual. San Antonio, TX: The Psychological Corporation; 1996.

20. Linehan MM, Comtois KA, Murray AM, Brown MZ, Gallop RJ, Heard HL, et al. Two-year randomized controlled trial and follow-up of dialectical behavior therapy vs therapy by experts for suicidal behaviors and borderline personality disorder. Arch Gen Psychiatry 2006 Jul;63(7):757-766. [doi: 10.1001/archpsyc.63.7.757] [Medline: 16818865]

21. Hawton K, Townsend E, Arensman E, Gunnell D, Hazell P, House A, et al. Psychosocial versus pharmacological treatments for deliberate self harm. Cochrane Database Syst Rev 2000(2):CD001764. [doi: 10.1002/14651858.CD001764] [Medline: 10796818]

22. Brown GK, Ten Have T, Henriques GR, Xie SX, Hollander JE, Beck AT. Cognitive therapy for the prevention of suicide attempts: a randomized controlled trial. JAMA 2005 Aug 3;294(5):563-570. [doi: 10.1001/jama.294.5.563] [Medline: 16077050]

23. Williams JM, Duggan DS, Crane C, Fennell MJ. Mindfulness-based cognitive therapy for prevention of recurrence of suicidal behavior. J Clin Psychol 2006 Feb;62(2):201-210. [doi: 10.1002/jclp.20223] [Medline: 16342287]

24. Beck A, Steer R. Manual for the Beck Hopelessness Scale. San Antinio, TX: The Psychological Corporation; 1988.

25.   Jung T, Wickrama KAS. An introduction to latent class growth analysis and growth mixture modeling. Social Pers Psych Compass 2008 Jan;2(1):302-317. [doi: 10.1111/j.1751-9004.2007.00054.x]

26.   Ram N, Grimm KJ. Growth mixture modeling: a method for identifying differences in longitudinal change among unobserved groups. Int J Behav Dev 2009;33(6):565-576 [FREE Full text] [doi: 10.1177/0165025409343765] [Medline: 23885133]

27.   Muthén L, Muthén B. Mplus Users Guide. Seventh edition. Los Angeles, CA: Muthén & Muthén; 2012.

28.   Asparouhov T, Muthén B. Auxiliary variables in mixture modeling: three-step approaches using M plus. Struct Equation Modeling 2014 Jun 09;21(3):329-341. [doi: 10.1080/10705511.2014.915181]

29.   van Spijker BA, Majo MC, Smit F, van SA, Kerkhof AJ. Reducing suicidal ideation: cost-effectiveness analysis of a randomized controlled trial of unguided web-based self-help. J Med Internet Res 2012;14(5):e141 [FREE Full text] [doi: 10.2196/jmir.1966] [Medline: 23103835]

30.   Hawton K, van Heeringen K. Suicide. Lancet 2009 Apr 18;373(9672):1372-1381. [doi: 10.1016/S0140-6736(09)60372-X] [Medline: 19376453]

Abbreviations

adj.BIC: sample size-adjusted Bayesian information criteria
AIC: Akaike information criteria
BDI: Beck Depression Inventory
BIC: Bayesian information criteria
BSS: Beck Scale for Suicide Ideation
LGMM: latent growth mixture modelling
OR: odds ratio
SI: suicidal ideation

XSL•FO
RenderX

<u>Original Paper</u>

Web Conversations About Complementary and Alternative Medicines and Cancer: Content and Sentiment Analysis

Mauro Mazzocut[1], MS(LIS); Ivana Truccolo[1], MSc; Marialuisa Antonini[2], MSc; Fabio Rinaldi[3], MSc Comp Sc, PhD; Paolo Omero[4], MSc(Comp Sci), PhD; Emanuela Ferrarin[1], MSc Pharm; Paolo De Paoli[5], M.D; Carlo Tasso[4], Eng.

[1]CRO Aviano National Cancer Institute, Scientific and Patient Library, Aviano, Italy

[2]Infofactory, Udine, Italy

[3]Institute of Computational Linguistics, University of Zurich, Zurich, Switzerland

[4]Artificial Intelligence Laboratory, Department of Mathematics, Computer and Physical, University of Udine, Udine, Italy

[5]CRO Aviano National Cancer Institute, Scientific Directorate, Aviano, Italy

Corresponding Author:
Mauro Mazzocut, MS(LIS)
CRO Aviano National Cancer Institute
Scientific and Patient Library
via Franco Gallini 2
Aviano, 33081
Italy
Phone: 39 0434659709
Fax: 39 0434659358
Email: mmazzocut@cro.it

Abstract

Background: The use of complementary and alternative medicine (CAM) among cancer patients is widespread and mostly self-administrated. Today, one of the most relevant topics is the nondisclosure of CAM use to doctors. This general lack of communication exposes patients to dangerous behaviors and to less reliable information channels, such as the Web. The Italian context scarcely differs from this trend. Today, we are able to mine and analyze systematically the unstructured information available in the Web, to get an insight of people's opinions, beliefs, and rumors concerning health topics.

Objective: Our aim was to analyze Italian Web conversations about CAM, identifying the most relevant Web sources, therapies, and diseases and measure the related sentiment.

Methods: Data have been collected using the Web Intelligence tool ifMONITOR. The workflow consisted of 6 phases: (1) eligibility criteria definition for the ifMONITOR search profile; (2) creation of a CAM terminology database; (3) generic Web search and automatic filtering, the results have been manually revised to refine the search profile, and stored in the ifMONITOR database; (4) automatic classification using the CAM database terms; (5) selection of the final sample and manual sentiment analysis using a 1-5 score range; (6) manual indexing of the Web sources and CAM therapies type retrieved. Descriptive univariate statistics were computed for each item: absolute frequency, percentage, central tendency (mean sentiment score [MSS]), and variability (standard variation σ).

Results: Overall, 212 Web sources, 423 Web documents, and 868 opinions have been retrieved. The overall sentiment measured tends to a good score (3.6 of 5). Quite a high polarization in the opinions of the conversation partaking emerged from standard variation analysis (σ≥1). In total, 126 of 212 (59.4%) Web sources retrieved were nonhealth-related. Facebook (89; 21%) and Yahoo Answers (41; 9.7%) were the most relevant. In total, 94 CAM therapies have been retrieved. Most belong to the "biologically based therapies or nutrition" category: 339 of 868 opinions (39.1%), showing an MSS of 3.9 (σ=0.83). Within nutrition, "diets" collected 154 opinions (18.4%) with an MSS of 3.8 (σ=0.87); "food as CAM" overall collected 112 opinions (12.8%) with a MSS of 4 (σ=0.68). Excluding diets and food, the most discussed CAM therapy is the controversial Italian "Di Bella multitherapy" with 102 opinions (11.8%) with an MSS of 3.4 (σ=1.21). Breast cancer was the most mentioned disease: 81 opinions of 868.

Conclusions: Conversations about CAM and cancer are ubiquitous. There is a great concern about the biologically based therapies, perceived as harmless and useful, under-rating all risks related to dangerous interactions or malnutrition. Our results can be useful to doctors to be aware of the implications of these beliefs for the clinical practice. Web conversation exploitation could be a strategy to gain insights of people's perspective for other controversial topics.

XSL•FO
RenderX

Introduction

Background

Complementary and alternative medicine (CAM) is an umbrella definition for a wide range of medical practices that are no part of the standard medical care [1,2]. These therapies are considered "complementary" if used in addition to a conventional treatment and "alternative" when used instead of it [3].

A survey of the National Institutes of Health reports an increasing trend in CAM usage among American people from 1999-2012 [4]. This trend can be easily compared with similar data that emerged in Europe [5], Asia [6,7], and Africa [8]. The scientific literature about CAM is also growing constantly [5,9-11]. Apart from the research on clinical effectiveness of CAM, many studies on patients' behaviors and choices are available [11-16]. Today, one of the most relevant topics in literature is the nondisclosure of CAM use among cancer patients [9,13,17].

The Italian context scarcely differs from these trends. A national survey performed by Italian Center for Social Studies and Investments (CENSIS) reports that 23.4% of Italian people resorted to unconventional medicine such as homeopathy and herbal remedies at least once in 2008 [18]. D'Arena [14] reported that CAM use among Italian cancer patients is expanding and mostly self-administered.

Current literature reports that patients would like to talk with their doctors about CAM, but they do not [5,6,9,16]. In particular, they would like to ask for information about safe use of CAM, rather than scientific evidence [16,19]. Today, many studies report that this information need is often unmet [16,17,20,21].

On one hand, usually doctors do not ask patients about their CAM consumption habits [9]. Maybe this is because of doctor's lack of knowledge of CAM issues [9,15,19,22]. On the other hand, patients anticipate doctors' disapproval or consider the disclosure irrelevant to their conventional care [9,23]. Consequently, patients autonomously search for information about CAM.

This general lack of communication between doctors and patients about CAM has relevant consequences.

First, it exposes patients to potentially dangerous behaviors [9,14,15,23,24]. Second, health professionals miss relevant information about patients' needs, beliefs, behaviors, and experiences about CAM therapies [6,13,25-27]. Third, lack of communication eases the spread of misinformation [21,28]. False information has relevant effects not only on patients' health outcomes [29,30] but also on the decision makers' policy as well, through collective debates [28,31,32]. Fourth, the lack of communication forces people to use alternative and less-reliable information channels [20,33-36]. Several studies reported that social communities are the main alternative information source about CAM for patients [5,37,38]. Within them, health information flows as a word of mouth mainly driven by narration of people's experiences, emotions, and opinions [5,28,37,39], in an intense interaction with mainstream media [31,32].

This trend is amplified through the virtual communities that crowd the Web [39,40]. Today, we are able to mine and analyze systematically the amount of unstructured textual information available in websites, forum, social networks, and other digital communities [41]. Some studies extracted relevant information about patients' point of view about different health topics exploiting social network analysis, natural language processing, content, and opinion mining software [42-44]. This knowledge is implicit in the Web conversations, in the semantic relationships among users, and in their opinions expressed by tags and comments. We hypothesize that the Web is rich of information concerning people's perspective about CAM topics. Even if expressed in an informal context, Web conversation can give an insight about the untold Italian people opinions, beliefs, and rumors about CAM. This is an implicit knowledge otherwise achievable through the Scientific and Patients Library (SPL) interviews or formal questionnaires.

Preliminary Data

We collected data about CAM information requests to the SPL of the Centro di Riferimento Oncologico (CRO) Aviano, National Cancer Institute. From 2008 to 2013, 218 of 2313 overall questions requested to SPL (9.4%) were CAM-related. But the trend is constantly increasing. In 2008, we recorded 24 CAM-related questions of 387 (6.2%). In 2012, the CAM-related questions were 38 of 198 (19.2%), and in 2013, 46 of 282 (16.3%). Overall, 183 requests of 218 (84%) regarded biologically based therapies such as diet, food, or natural remedies effectiveness; interactions among remedies and drugs; and remedies availability. We have to point out that these numbers reflect only the questions recorded through the SPL access form. Due to our policy [45], this form is filled by patients to respect their privacy. Consequently, they are not forced to give any information about their requests.

Aims

Based on the literature and the preliminary data, we aim to analyze Italian Web conversation about CAM. In particular, we aim to identify the most relevant Italian Web conversation sources; the cancer CAM therapies most discussed by Italian Web users and the corresponding sentiment; and what is the most discussed disease.

Methods

Our data collection methodology could be assimilated to the theoretical sampling [46] because it involved simultaneously collecting, coding, and analyzing textual data from the Web conversations. The textual data analyzed in this work have been collected by means of ifMONITOR [47], an automatic web intelligence tool developed by the Artificial Intelligence Laboratory of the University of Udine. ifMONITOR provided the following services: (1) Web filtering, (2) Web monitoring and thematic database construction, and (3) automatic classification.

Within our search, we followed a systematic approach, which was organized into 6 phases, with a strictly functional approach. We have used ifMONITOR for services "a" and "b" in phase 3 and for service "c" in phase 4. Figure 1 shows the data collection and analysis flowchart. Further details and a flowchart about the complete workflow are available in the Multimedia Appendix 1.

Figure 1. Web conversation analysis flowchart.
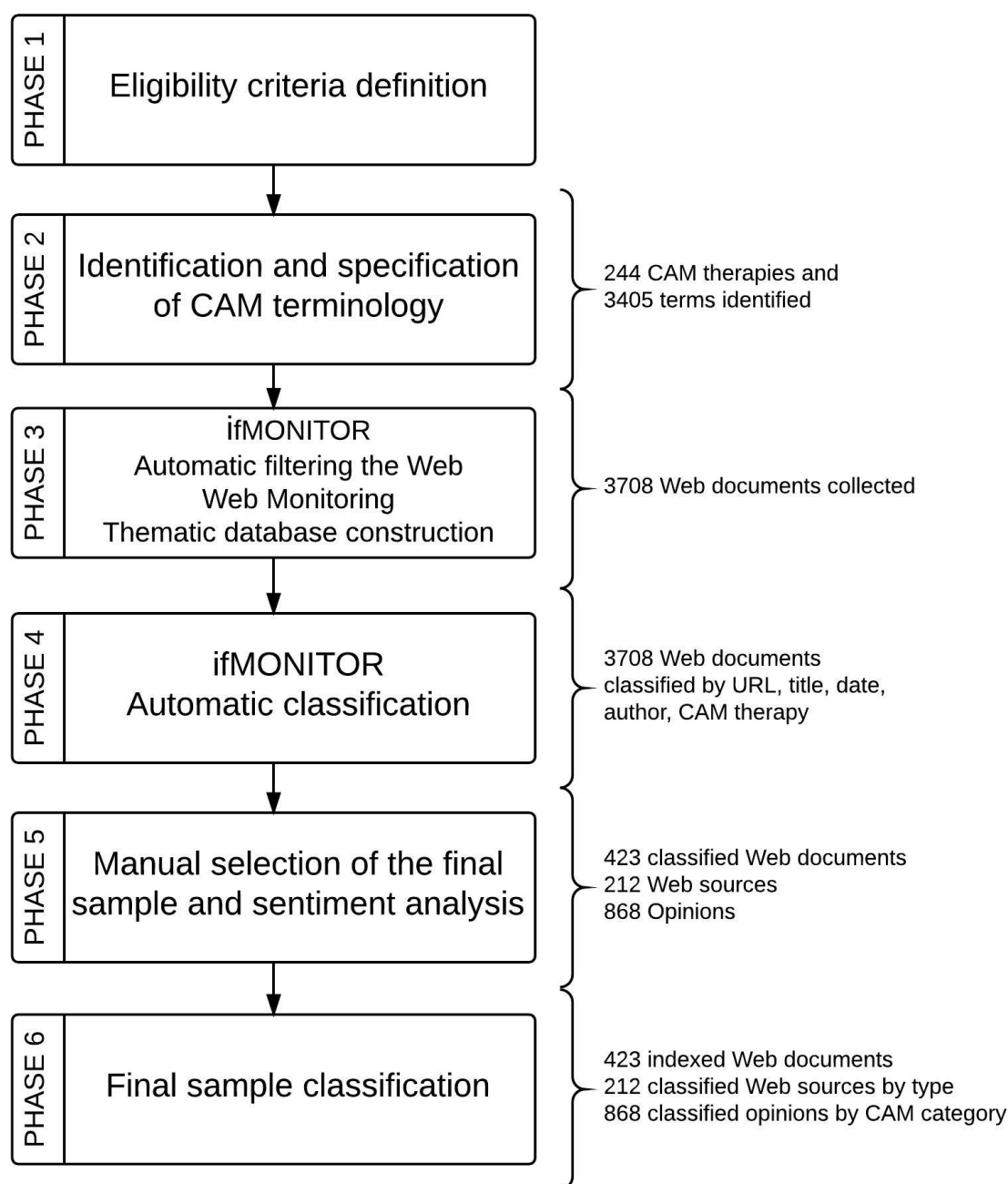
Phase 1: Eligibility Criteria Definition

Websites, forums, blogs, communities, and social networks of any kind and topic addressed to the general public and containing conversations about CAM topics have been included. In particular, Facebook profiles, pages, and groups of users and YouTube channels were also included (depending on the privacy settings). Otherwise, we decided to leave out Twitter because the length of the messages is too short to express a complex opinion about the topic taken into account. The Web search was also limited to Web documents published between January 1, 2013, and May 31, 2014. We also limited the Web search only to Italian language Web documents, to meet the perspective of Italian people with low literacy skills.

Phase 2: Identification and Specification of CAM Terminology

This phase is manual and has been executed by CRO librarians. They built up a CAM database containing relevant terms to be used as a reference in subsequent phases. For each CAM therapy identified, a structured template was filled up with related synonymies and references to specific philosophy or religion; names of people, organizations, and places; book titles; anatomy and physiology terms; substances or drugs; principles or processes; methods; instruments or tools; effects or side effects; diseases. Because the final goal was to analyze Web-based people conversations, it was important to capture the terminology that is actually known by the people themselves. For these reasons, 3 main kinds of sources of information have been considered: (1) scientific and medical resources [1,48,49], (2) specific educational booklets for patients published by cancer volunteer associations [50,51], and (3) public Web-based sources concerning CAM in Italian language easily accessible on the Web [52,53]. Figure 2 shows a sample of our CAM database referring to the "Gerson Therapy".

Figure 2. CAM database sample: "Gerson Therapy" item translated into English.

CAM	Gerson Therapy
SYNONIMIS	Gerson diet
PHYLOSOPHY,	-
PEOPLE	Max Gerson, Charlotte Gerson, Claude Pepper
BODIES	Gerson Institute
BOOKS	A Cancer Therapy: Results of Fifty Cases
ANATOMY, PHYSIOLOGY	Cellular metabolism, bile
DRUGS OR SUBSTANCES	Fruit Juice, Fruit and Row Vegetables, Coffee, Thyroid Gland Extract, Phosphoric Acid, Lugol's Iodine, Potassium, Pancreatin, Vitamin C
PRINCIPLES OR	Digestion, Acidity, Homeostasis, Alkalinity, Sodium Potassium Imbalance
TECHNIQUES	Diet, Vegetarian Diet, Dietary Supplements Enemas of Coffee, Bile Stimulation
TOOLS	-
EFFECTS	Detox, Self Healing, Toxine Removal. Cancer Regression, Antineoplastic Effect
DISEASES	Cancer
PLACES	Wongrowitz, Albert Ludwigs Universitat, Freiburg, Bielefield, Tijuana

Phase 3: Automatic Filtering of the Web

This phase is aimed at collecting all the materials to be analyzed later exploiting ifMONITOR services "a" and "b."

The final result of phase 3 has been a set of 3708 documents, largely referring to people's conversations, which were processed in phase 4.

Phase 4: Automatic Classification

This fully automated phase is aimed at classifying the 3708 documents retrieved in which CAM therapy is mentioned in the document. The starting point has been the terms included in the CAM therapies database.

Phase 5: Selection of the Final Sample and Sentiment Analysis

This phase was mainly manual, and it was aimed at identifying a sample of documents to be analyzed in detail. The result constituted 838 Web documents.

The manual analysis of the documents, starting from the documents of the most successful Web-based sources (websites, blog, and forums that published a high number of pages concerning CAM therapies). These criteria enabled us to consider the most popular sources and Web conversations first.

This process continued up to the accumulation of 423 Web documents, when we reached the saturation of the sample [46]. We considered this a significant and feasible size for the sample of documents to be manually analyzed in detail because the rest of the Web documents retrieved did not shed any further light to the evaluation on the most relevant Web conversation sources or the most mentioned CAM therapies.

For each of the 423 documents, manual tagging was added, concerning the sentiment about the perceived effectiveness or ineffectiveness of the CAM therapy mentioned through a numerical score ranging from 1 to 5 (1 very bad; 2 bad; 3 neutral; 4 good; and 5 very good). The score assignment was performed by a team of 4 trained Web intelligence analysts. They performed individually the sentiment analysis on different sets of documents, later they compared together the score assignment, thus limiting information bias due to interoperator discretionary.

We did not perform any assessment of the medical and scientific reliability of these statements. Figure 3 shows a screenshot of ifMONITOR thematic database.

Figure 3. ifMONITOR thematic database screenshot. Highlighted in yellow are the opinions stated in the Web document. In capitals is the number of relevant occurrences automatically indexed by ifMONITOR: allopathic medicine (chemotherapy and radiotherapy); Di Bella Multitherapy, acupuncture. In italics followed by a red, green, or gray spot, the sentiment analysis about the issues stated in the document. Red spot (bad sentiment): plot against Di Bella Multitherapy; allopathic medicine carcinogenic effects. Green spot (good sentiment): acupuncture effectiveness; Di Bella Multitherapy effectiveness; Cannabis as palliative care effectiveness. Gray spot (neutral sentiment): melatonin in Di Bella Multitherapy; Uma de Gato effectiveness; noni and xeronina effectiveness.

Phase 6: Final Sample Classification

According to the literature [1,5,24,49], the librarians tagged the documents included in the final sample using five CAM therapies classes. Within each class, they grouped CAM therapies into 14 categories as shown in Table 1.

Furthermore, librarians indexed the Web sources according to the criteria of health and nonhealth thematic areas. Web sources are websites, blogs, forums, and social networks (personal profiles, groups of users, or public pages). Then they gathered Web sources according to their specific subject matters, as shown in Table 2.

The content classification allowed us to identify the most discussed CAM therapies; the most relevant Web sources; and

the most discussed cancer type. Furthermore, the semantic analysis allowed us to measure the popularity and sentiment about the identified CAM and identify the pros and cons issues in the discussions about the most discussed CAM therapies.

Statistical Analysis

Descriptive statistical analysis was performed using Microsoft Excel. Absolute (n) and relative frequencies (%) with reference to total Web sources, total Web documents, and/or total opinions included in the final sample were computed. For each item expressed using a numerical score (Likert-type scale ranging from 1=very bad to 5=very good) mean scores (mean sentiment score [MSS]) as a central tendency index and standard deviations from the mean (σ) as an indicator of variability were also calculated.

Table 1. CAM therapies.

CAM category		Description
Biologically-based		
	Nutrition	Food or diets (ie, veganism, Gerson diet, garlic, green tea)
	Chemical-based medicines	Drugs or remedies based on chemical substances (ie, UK101, baking soda)
	Plant-based medicines	Drugs or remedies derived from plants (ie, Aloe, Essiac, Laetrile)
	Animal-based medicines	Drugs or treatments derived from animals (ie, Escozul, Shark fin)
	Mineral-based medicines	Drugs or treatments derived from minerals (ie, Alcaline water)
Whole medical systems		
	Alternative medical systems	Medical systems apart from standard medicines (ie, Olism, New German Medicine, Omeopathy)
	Traditional medical systems	Remedies and treatments based on folk traditions (ie, traditional Chinese medicine, Ayurvedic medicine)
Mind body interventions		
	Meditation and relaxation	Techniques exploiting the mind's ability to affect biological functions (ie, mindfulness, yoga, hypnosis)
	Psychological approaches	Psychological approaches or stimulation to release the stress (ie, Simonton method, music therapy, color therapy)
	Spirituality	Techniques based on supernatural or divine intervention (ie, pray, self-healing, faith)
Energy therapies		
	Natural energies	Techniques based on natural energies such as heat (ie, hyperthermia, bioenergy, hydrotherapy)
	Spiritual energies	Techniques based on supernatural energies (ie, pranotherapy, Reiki)
Manipulative and body-based		
	Exercises	Techniques based on physical exercises (ie, dance therapy, Tai Chi, silvotherapy)
	Massages	Techniques based on body manipulations (ie, reflexology, osteopaty, chiropratic)

Table 2. Web sources classification

Web sources category		Description
Health-related		
	CAM	Fostering of CAM therapies, including cancer CAM therapies
	Health and wellness	Health communication and healthy lifestyles promotion
	Oncology	Cancer patients communities
Nonhealth-related		
	Mainstream newspaper on the Web	Newspaper, both national or local, print or Web-based
	Pseudoscience	Fostering of pseudoscientific or plot theories
	Debunking	Exposition of Web-based pseudoscientific theories
	Other	Any other issues retrieved

Results

Phase 2: Identification of CAM Terminology

The CAM Therapy Database included a set of 224 CAM therapies: 19 cancer treatments (ie, Essiac, Gerson Therapy, Escoazul); 205 general remedies used to manage diseases symptoms, treatments side effects, or to cope with the stress (ie, meditation). The number of terms collected referring to the 224 CAM therapies is 3405.

Phase 3 to 6: From Classification to Sentiment Analysis

XSL·FO
RenderX

Web Documents Overall

The final sample included 423 Web documents extracted from 212 Web sources, showing 868 opinions.

Ten Web sources including 39 Web documents tagged automatically by ifMONITOR have become inaccessible before Phase 6. Nevertheless, all of them were automatically indexed by URL, title, author, date, CAM therapy mentioned, and opinion. But they lack information about Web sources or Web documents subjects (for instance topic of the single Web document). The 868 opinions retrieved stated an overall MSS of 3.6 (σ=1.0).

We found 244 of 868 (28.1%) opinions related to a specific disease: 81 of 244 (33.2%) were breast cancer related; 64 (26.2%) were pancreatic cancer related; 29 (11.9%) were melanoma related; 27 (11.1%) were lung cancer related; 21 (8.6%) were bone marrow cancer related; 14 (5.7%) were ovarian cancer related; 6 (2.5%) were peritoneum cancer related; 1 (0.4%) were neck cancer; and 1 (0.4%) were mandible cancer related.

Web Sources

We found 95 of 212 websites (44.8%), 45 social networks (21.3%), 41 blog (19.3%), and 28 forums (13.2%).

Within social networks, 22 of 45 (49%) were public personal profiles, 12 (27%) were groups of users, and 11 (24%) were company or personal pages.

We found 83 of 212 (39.2%) "health-related" Web sources producing 176 of 423 (41.6%) Web documents. Among the "health-related" group, we found 46 of 212 (21.7%) Web sources focused on CAM topics producing 103 of 423 (24.3%) Web documents. The 377 opinions of 868 (43.4%) of this group, expressed a MSS of 3.7 (σ=0.8). The Web sources focused on "oncology" topics were 8 of 212 (3.8%), producing 37 of 423 (8.7%) Web documents.

On the other hand, 126 of 212 (59.4%) were "nonhealth-related" Web sources, producing 237 of 423 (56%) Web documents. Among "nonhealth-related" group, we found 66 of 212 (31.1%) Web sources focused on a wide range of different topics that produced 142 of 423 Web documents (33.6%). The 476 opinions of 868 (54.8%) retrieved into this group, expressed a MSS of 3.6 (σ=1.1).

The highest MSS (3.7) was observed among "other issues," "oncology," "health and wellness," and "CAM" categories. The lower (2.7) within the "debunking" category. Complete results are displayed in Table 3.

Table 3. Results per Web source.

Web Sources category		WS[a], n (%)	WD[b], n (%)	OP[c], n (%)	MSS[d] ($\pm\sigma$[e]).
Health-related		83 (39.2)	176 (41.6)	377 (43.4)	3.7 (± 0.8)
	CAM	46 (21.7)	103 (24.3)	198 (22.8)	3.7 (±0.8)
	Health and wellness	29 (13.7)	36 (8.5)	79 (9.1)	3.7 (±0.1)
	Oncology	8 (3.8)	37 (8.7)	100 (11.5)	3.7 (±1.1)
Nonhealth-related		126 (59.4)	237 (56)	476 (54.8)	3.6 (±1.1)
	Mainstream newspaper on the Web	26 (12.3)	35 (8.3)	93 (10.7)	3.4 (±1.1)
	Pseudoscience	30 (14.2)	52 (12.3)	124 (14.3)	3.6 (±1.1)
	Debunking	4 (1.9)	8 (1.9)	13 (1.5)	2.7 (±1.2)
	Other	66 (31.1)	142 (33.6)	246 (28.3)	3.7 (±1.1)
Unknown		3 (1.4)	10 (2.4)	15 (1.7)	3.9 (±0.3)
Total		212 (100)	423 (100)	868 (100)	3.6 (±1.0)

[a]WS: Web sources.

[b]WD: Web documents.

[c]OP: opinions.

[d]MSS: mean sentiment score.

[e]σ: standard deviation.

Yahoo Answers is the most retrieved Web source: 41 of 423 (9.7%). The Facebook group "Quelli che il cancro (e non solo) lo curo a modo mio" (those who cure cancer, and not only, in their own way), counted 30 Web documents of 423 (7.1%). This group changed its privacy settings: today the access is limited to the registered members only. But if we sum all Facebook profiles, groups, and pages together, we count 89 of 423 (21%) Web documents, nearly double that of Yahoo Answers. The third Web source per number, 14 of 423 (3.3%), was Greenstyle.it, a Web magazine fostering ecologic lifestyles owned by HTML.it, an Italian network of publishing and advertising companies.

CAM Therapies

We found 94 therapies or remedies mentioned in the conversations retrieved. Among these, 68 (72%) were included

in the CAM keyword set. Twenty-six (28%) are therapies that emerged during the manual reviews of ifMONITOR's search output in phase 3. Some of these are typical Italian alternative cancer remedies, such as "Giuseppe Nacci's" diet, an Italian variation of "Gerson Therapy," "Gianfranco Pantellini's" therapy, based on potassium ascorbate consumption, "Alessiani's Water," enriched with minerals extracted from Romans catacombs.

The most discussed CAM therapies belong to the "biologically based therapies" group, with 702 of 868 (80.9%) opinions retrieved. The MSS observed was 3.6 (σ=1.0). Among "biologically based therapies" the "nutrition" category had the higher rate of opinions, 339 of 868 (39.1%), expressing an MSS of 3.9 (σ=0.8). "Chemical-based medicines," counted 186 opinions of 868 (21.4%) showing an MSS of 3.3 (σ=1.3). "Plant-based medicines" counted 144 opinions of 868 (16.6%), and an MSS of 3.6 (σ=1.0). The complete results are displayed in Table 4.

Table 4. CAM therapies per category.

CAM category		OP[a], n (%)	MSS[b], ($\pm\sigma$[c])
Biologically-based therapies		702 (80.9)	3.6 ($\pm$1.0)
	Nutrition	339 (39.1)	3.9 ($\pm$0.8)
	Chemical-based medicines	186 (21.4)	3.3 ($\pm$1.3)
	Plant-based medicines	144 (16.6) 3.6 ($\pm$1.0)	
	Animal-based medicines	26 (3.0)	3.5 ($\pm$0.6)
	Mineral-based medicines	7 (0.8)	4.0 ($\pm$0.6)
Whole medical systems		103 (11.9)	3.5 ($\pm$1.0)
	Alternative medical systems	81 (9.3)	3.4 ($\pm$1.1)
	Traditional medical systems	22 (2.5)	4.0 ($\pm$0.8)
Mind–body interventions		32 (3.7)	4.0 ($\pm$0.6)
	Meditation and relax	20 (2.3)	4.1 ($\pm$0.7)
	Psychological approaches	6 (0.7)	4.2 ($\pm$0.4)
	Spirituality	6 (0.7)	3.8 ($\pm$0.4)
Energy therapies		20 (2.3)	3.9 ($\pm$0.7)
	Natural energies	15 (1.7)	4.0 ($\pm$0.6)
Spiritual energies	5 (0.6)	3.4 ($\pm$0.9)	
Manipulative and body-based		11 (1.3)	3.8 ($\pm$0.4)
	Exercises	8 (0.9)	3.9 ($\pm$0.3)
	Massage techniques	3 (0.3)	3.7 ($\pm$0.6)
Total		868 (100)	3.6 ($\pm$1.0)

[a]OP: opinions.

[b]MSS: mean sentiment score.

[c]σ: standard deviation.

We identified different approaches concerning the "nutrition" issues. On one hand, most of the time, Web users talked about the effectiveness of a certain diet or alimentary regimen in cancer treatment or prevention. For instance, "vegetarianism" or "veganism," "raw foodism," "alkaline diet," "Gerson therapy." On the other hand, they often discussed about the healing properties of a specific food or spice. For instance, grapefruit, garlic, turmeric, mushrooms, lemon, or green tea.

Finally, the discussions on dietary regimen included also the use of supplements, for instance, vitamins. "Diets" overall collected 154 opinions of 868 (18.4%) with an MSS of 3.8 (σ=0.9). "Food as CAM" overall collected 112 opinions of 868 (12.8%) with a MSS of 4 (σ=0.7). "Supplements" collected 73 opinions of 868 (7.8%) with an MSS of 3.8 (σ=0.9). The main results are summarized in Table 5.

XSL•FO

RenderX

Excluding diets and food, the most discussed CAM therapy is the "Di Bella multitherapy." We retrieved 102 opinions of 868 (11.8%) with an MSS of 3.4 (σ=1.2). "Di Bella multitherapy" and a reportage of "Le Iene,"(the Hyenas) Italian comedy and

satirical TV show with reports on politics and consumers issues, on "alkaline diet" and "veganism" were also the most recurring topic: 16 of 423 (3.8%) Web documents retrieved.

Table 5. Most discussed CAM therapies.

CAM classification		CAM category	OP[a], n (%)	MSS[b], (±σ[c])
Diets				
	Biologically-based therapies Nutrition	154 (18.4)	3.8 (±0.9)	
Food as CAM				
	Biologically-based therapies	Nutrition	112 (12.8)	4.0 (±0.7)
Di Bella multitherapy				
	Biologically-based therapies	Chemical-based medicines	102 (11.8)	3.4 (±1.2)
Phytotherapy				
	Biologically-based therapies	Plant-based medicines	96 (11.1)	3.6 (±1.0)
Supplements				
	Biologically-based therapies	Nutrition	73 (7.8)	3.8 (±0.9)
Simoncini therapy				
	Biologically-based therapies	Chemical-based medicines	45 (5.2)	2.6 (±1.4)
Naturopathy				
Whole medical systems	Alternative medical systems	38 (3.7)	3.5 (±1.0)	
Father Zago's aloe remedy				
	Biologically-based therapies	Plant-based medicines	22 (2.5)	3.6 (±0.9)
Others				
		-	226 (26)	-

[a]OP: opinions.

[b]MSS: mean sentiment score.

[c]σ: standard deviation.

The Case of Di Bella Multitherapy

We found 29 of 212 (13.7%) Web sources containing 58 of 423 (13.7%) Web documents and 102 of 868 (11.7%) opinions about "Di Bella multitherapy." This is a controversial cancer treatment invented by the Italian physician Luigi Di Bella, based on a mix of somatostatin, melatonin, hormones, and vitamins. In the early 1990s, this treatment had a wide appeal among Italian cancer patients [54]. The trial performed in 1998 by the Italian Ministry of Health proved that this therapy is ineffective [54-58]; nevertheless, it has been followed by an intense emotional campaign in favor of the therapy [54]. For these reasons, patients rarely ask for information or disclose the use of Di Bella multitherapy to their doctors. Three Web sources containing 8 Web documents were not accessible in phase 6. The complete results are displayed in Table 6.

XSL•FO

RenderX

Table 6. Discussions about Di Bella multitherapy.

CAM category		WS[a], n (%)	WD[b], n (%)	OP[c], n (%)	MSS[d], (σ[e])
Health-related		8 (28)	15 (26)	37 (36.3)	3.4 ($\pm$1.1)
	CAM	3 (10)	4 (7)	8 (7.9)	3.0 ($\pm$1.1)
	Health and wellness	0 (0)	0 (0)	0 (0)	-
	Oncology	5 (17)	11 (19)	29 (28.4)	3.6 ($\pm$1.1)
Nonhealth-related		21 (72)	43 (74)	65 (63.7)	3.3 ($\pm$1.3)
	Mainstream newspaper on the Web	6 (21)	9 (15)	13 (12.7)	3.5 ($\pm$1.2)
	Pseudoscience	6 (21)	9 (15)	17 (16.6)	3.8 ($\pm$1.1)
	Debunking	0 (0)	0 (0)	0 (0)	-
	Other	9 (31)	25 (43)	35 (34.4)	3.1 ($\pm$1.3)
Total		29 (100)	58 (100)	102 (100)	3.4 ($\pm$1.2)

[a]WS: Web sources.

[b]WD: Web documents.

[c]OP: opinions.

[d]MSS: mean sentiment score.

[e]σ: standard deviation.

This therapy was mostly discussed among "Nonhealth-related" web sources, with 65 out of 102 (63.7%) opinions retrieved showing an MSS of 3.4 (σ=1.1). Most of the conversation were retrieved among the "other issues" category, with 35 of 102 (34.4%) opinions showing an MSS of 3.1 (σ=1.1). No opinions were retrieved into the "debunking" and "health and wellness" category.

Conversely, we counted 37 of 102 (36.3%) opinions into "health-related" Web sources showing an overall MSS of 3.3 (σ=0.8). We observed that it was discussed almost only in the "oncology" category, with 29 of 102 (28.4%) opinions retrieved showing an MSS of 3.6 (σ=1.1).

We also found that 23 of 423 (5.4%) Web documents have the "Di Bella multitherapy" as main topic of the conversation, distributed among: 6 in "oncology," 1 in "CAM," 2 in "mainstream newspaper on the Web," 2 in "pseudoscience and conspiracy theories," and 12 in "other issues" categories.

The conversations retrieved were mainly focused on the discussion about effectiveness or ineffectiveness of "Di Bella therapy," with 60 of 102 (58.8%) opinions showing an MSS of 3.8 (σ=1.3). The lowest MSS observed is 2.5 (σ=1.1) achieved by the topic "plot against therapy's trial" means that the 5 users believe to the plot theory against the "Di Bella multitherapy."

The arguments in support of the Di Bella multitherapy expressed by users in the Web conversations were the absence of side effects, the effectiveness in treating neoplasm otherwise terminal, and its effectiveness while conventional therapies are suspended. There is a widespread belief that its real effectiveness cannot be proven because the trial was intentionally compromised. To prove this last statement, users often report that the well-known oncologist Umberto Veronesi during an interview confirmed the effectiveness of the somatostatin in breast cancer treatment. The points against the Di Bella multitherapy were mainly two: the therapy has not been

recognized as a standard therapy by the Italian National Health Service because of the lack of scientific evidence and the therapy has high initial costs.

Discussion

Principal Findings

Our survey confirms the existence of Web-based conversations about cancer CAM therapies among Italian Web users. The overall sentiment on CAM effectiveness tends to range from neutral to good, with a certain degree of variability in opinions (MSS$\pm$: 3.6$\pm$1.0).The conversations about CAM were retrieved mostly from "nonhealth-related" rather than "health-related" web sources. Facebook and Yahoo Answers together covered almost one third of the conversation retrieved. Beyond them, most of the CAM conversations are widespread within Web sources that deal with very different topics. For instance, political activism, economy, marketing, video games, sport, gastronomy, leisure, and weather forecasts. The overall sentiment observed in the "nonhealth-related" category is quite neutral with high level of polarization (MSS$\pm$: 3.3$\pm$1.3).

A similar overall sentiment emerges in the "health-related" category, although with less variability (MSS$\pm$: 3.4$\pm$1.1). Despite the low number of Web sources retrieved, we found a high rate of opinions in the "oncology" category. This is because they are mostly conversations among patient-dedicated forums users. The sentiment observed in this category is quite good (MSS=3.8), with some variability among the conversations partaking (σ=1.1).

The most discussed CAM therapies belong largely to the "biologically-based" category. In particular, the "nutrition" has a good score with low polarization (MSS$\pm$: 3.9$\pm$0.8). "Nutritional" CAMs are also considered very effective for cancer prevention, side-effects management, and as cancer cure also. Nutritional conversations concern the effectiveness of

diets, such as "veganism." But they also concern the healing properties of a single food. In this case, the consensus is high and with quite low variation (MSS±: 44±0.9). "Plant-based" medicines are also perceived as quite effective (MSS±: 3.6±1.0). Against the number of opinion retrieved, we observed an almost neutral sentiment score with a quite high polarization of opinions (MSS±: 3.3±1.3) among "chemical-based" medicines.

Only a third (28.1%) of the opinions retrieved mentioned a specific neoplasm: mostly breast or ovarian cancer. In the other case, people referred only to cancer in general.

Contrary to what we expected, we collected a low number of very popular CAM, such as homeopathy, Chinese traditional medicine, acupuncture, yoga, and reflexology. At the same time, almost a third (28%) of the CAMs that emerged during the manual reviews in phase 3 were unexpected or uncommon. They are all "biologically based."

Finally, particularly significant is the finding of a considerable amount of opinions about the "Di Bella multitherapy." We observed that Di Bella multitherapy is mentioned mostly within the "nonhealth-related" Web sources. The corresponding MSS tends to a neutral value, with a very high rate of polarization, especially among the "nonhealth-related" Web sources (σ 1.3). It is interesting to observe that the higher MSS is recorded among the "oncology" and "pseudoscience and conspiracy theories" categories together.

These findings on Di Bella Multitherapy are particularly relevant if we compare its popularity with other similar pseudoscientific cancer treatments, such as the "Simoncini Baking Soda" therapy. This latter is notably less discussed, with a very low MSS and a very high level of polarization (σ 1.4). This proves that Di Bella multitherapy is considered more than a pseudocure by those partaking in the Web conversations that were considered.

Comparison With Prior Works

Our findings in number and distributions of conversations confirm those reported in the Social Oncology Project for the United States [39]: cancer conversations are ubiquitous. The sentiment expressed in the considered Italian Web conversations are comparable with those observed in Israel [6] about the perception of effectiveness and safety of CAM (mean scores were equal to 3.4619 and 3.6589, respectively, using a similar 1-5 score range).

The distribution of the conversations complies with the results of the CENSIS national report [38]: only 13.9% of the Italian people that are used to share advices and experiences in forum or specific Web-based communities. Conversely, almost 90% of Italian people exploit search engines to find the health information they need. Very generic Web sources such as Facebook and Yahoo Answers are the main conversation triggers. Facebook users seem to be more active in promoting diets and healthy lifestyles (ie, naturopathy and ecology). But we also found several pages that promoted pseudomedicine practitioners or products (ie, Escozul). Yahoo Answers users behavior differ if they are "askers" or "answerers." The askers show a rather low knowledge about cancer topics and about the differences among its typologies. They ask very generic questions such as: "What is the name of the disease that make

you go bald?" or "What is the cancer cure?" or "Is there a cure for cancer alternative to chemotherapy?" The answerers are also very generic in their replies. But we found that the most active answerers are strong promoters of alternative remedies rather than proper informers (ie, the motto "Di Bella multitherapy" used as profile's avatar).

This complies with Quattrociocchi [32] and Bessi [28], who observed that in science/pseudoscience Facebook conversations consumers of alternative news are more focused on their contents, whereas scientific news consumers are more prone to comment on alternative news. Moreover, Mocanu [31] observes that those with strong preferences for alternative information sources, perhaps motivated by the will to avoid the manipulation employed by mainstream media controlled by the government are more susceptible to false information. According to Del Vicario [59], Facebook users tend to select and share content related to specific narratives and ignore the rest. This way, conversations often occur in "echo chambers" which cause reinforcement of confirmation bias, shaping users' commitment to a specific system of beliefs and fostering the spread of misinformation. Particularly, Bessi [28] observed also that unsubstantiated scientific claims and rumors reverberate over a period of time comparable with that of more verified information among the Italian Facebook community. This increases the polarization of those for and against conventional medicine. On the other hand, Glaeser [60] demonstrates that if people have strong previous convictions, then the corrections of false information can backfire by increasing polarization among those participating in discussion. Above all, this is seen to be true in the persistence of the same issues about Di Bella multitherapy, over a period of 20 years, as reported by Passalacqua [54].

The amount of conversations retrieved about nutritional and herbal remedies together comply with the trends in CAM consumption reported by current literature [5,6,15,35,61]. In particular, D'Arena [14] reported that Italian cancer patients resort mostly to aloe formulations, green tea, and supplements. But in our findings, a clear distinction between nutrition and supplements and plant-based medicines emerges. We observed a wide consensus on food and diets effectiveness to face cancer. Conversely, plant-based medicines and supplements benefits are more argued. Moreover, "chemical-based" medicines are more controversial. This complies with CENSIS report [18]: remedies that claim a "natural" background are perceived as more effective and safe than the "chemical-based" ones. Web users tend to consider "nutritional" and "natural" remedies harmless. Conversely, food chemically or industrially processed such as meat, fish, carbohydrates, complex sugars, and milk products are considered too adulterated and harmful.

Despite the wide consumption of some popular CAM therapies reported by the literature among the Italian cancer patients [5,18,62], we collected a very small number of conversations about these topics. Probably, this is due to the fact that acupuncture, phytotherapy, homeopathy, homotoxicology, anthroposophic medicine, ayurvedic medicine, and Chinese traditional medicine are accepted as medical acts by professional bodies of Italian physicians [63], but also to the fact that the

"nutritional" issues are currently mostly discussed in other Italian media, such as television (ie, "Le Iene" TV show).

Contrary to the trends observed by the Social Oncology Project [39], the volume of Italian Web conversation is matched with the impact of the disease. In fact, the Italian Association of Cancer Registries [64] reports that breast cancer was the most common incident cancer among women in 2011 (13%) and also the most prevalent neoplasm from 1992 to 2006 (522,235 cases). But the prevalence of gynecological cancer conversation could be related also to some predictors of CAM use and the Internet health information consumption. Molassiotis [5], Jong [16], Yun [59], and Hubner [15] agree that women with chronic disease, high education, and good socioeconomic status are more inclined to use CAM. At the same time, Kelly [65] reported that active health information seeking using a wide range of sources was predicted by female gender, aged 40-64 years, higher education. The trend among Italian people is the same [14,18,66]. Siliquini [66] reported also that women with chronic disease, aged 42-53 years, are more susceptible to negative modification in health behaviors induced by Web information.

Limitations

Our study has some limitations. The sample is limited in number, and it is not representative of all the Italian Web conversations in the period taken into account. Furthermore, it was not possible to acquire demographic information about Web users. Although we have included a consistent number of Web users, this may not be representative of the entire Italian population. The use of mean and standard deviation for summarizing the opinions collected using a Likert-type ordinal scale is controversial; however, given the adequate sample size of each group and the nearly normal distribution of our data, the parametric approach is considered acceptable [66].

Conclusions

Conversations on CAM and cancer are ubiquitous. Most of them flow through nonhealth-related Web sources. We need to know people's information background to better understand their "implicit" knowledge about CAM issues and cancer. We found that unsubstantiated beliefs such as Di Bella multitherapy last for many years, despite scientific evidence. This particularly happens with the cancer patients' Web-based communities that were retrieved. Understanding the reasons for this persistency against scientific evidence needs further research. There is a big concern about the biologically based therapies and remedies, in particular diet and nutrition and naturopathy issues. These are perceived as harmless and useful, under-rating all risks related to dangerous interactions or malnutrition.

Based on our data, we suggest that doctors should be aware of the implications of these beliefs for the clinical practice. At the same time, it is necessary to keep themselves up to date about the ever-changing CAM issues, to foster patients' information requests and avoid decisions based only on autonomous Web searches.

We can exploit Web conversation analysis as a strategy to get insights of people perspective and improve new information strategies and resources even for other controversial topics such as clinical trials, sexuality and fertility, and HPV vaccination.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Extended description of the data collection and analysis workflow.

[PDF File (Adobe PDF File), 371KB - jmir_v18i6e120_app1.pdf]

References

1. National Center for ComplementaryIntegrative Health. Complementary, Alternative, or Integrative Health: What's In a Name Internet URL: https://nccih.nih.gov/health/whatiscam [accessed 2015-03-10] [WebCite Cache ID 6Wv8F0uzH]
2. National Cancer Institute.: National Cancer Institute Complementary and Alternative Medicine Internet URL: http://www.cancer.gov/cancertopics/cam [accessed 2015-03-10] [WebCite Cache ID 6Wv9Kvnd9]
3. National Library of Medicine, National Institute of Health. Medical Subject Headings. Complementary Therapies Internet URL: http://www.ncbi.nlm.nih.gov/mesh/68000529 [accessed 2015-03-10] [WebCite Cache ID 6Wv8Q0xhj]
4. National Center for Complementary Integrative Health. National Health Interview Survey 2012 URL: https://nccih.nih.gov/research/statistics/NHIS/2012 [accessed 2015-03-10] [WebCite Cache ID 6Wv8KD1gT]

5.   Molassiotis A, Fernadez-Ortega P, Pud D, Ozden G, Scott JA, Panteli V, et al. Use of complementary and alternative medicine in cancer patients: a European survey. Ann Oncol 2005 Apr;16(4):655-663 [FREE Full text] [doi: 10.1093/annonc/mdi110] [Medline: 15699021]

6.   Ben-Arye E, Frenkel M, Klein A, Scharf M. Attitudes toward integration of complementary and alternative medicine in primary care: Perspectives of patients, physicians and complementary practitioners. Patient Educ Couns 2008;70(3):402. [doi: 10.1016/j.pec.2007.11.019]

7.   Tsuya A, Sugawara Y, Tanaka A, Narimatsu H. Do Cancer Patients Tweet? Examining the Twitter Use of Cancer Patients in Japan. J Med Internet Res Internet 2014:27. [doi: 10.2196/jmir.3298]

8.   Sawadogo WR, Schumacher M, Teiten M, Dicato M, Diederich M. Traditional West African pharmacopeia, plants and derived compounds for cancer therapy. Biochem Pharmacol 2012 Nov 15;84(10):1225-1240. [doi: 10.1016/j.bcp.2012.07.021] [Medline: 22846603]

9.   Davis EL, Oh B, Butow PN, Mullan BA, Clarke S. Cancer patient disclosure and patient-doctor communication of complementary and alternative medicine use: a systematic review. Oncologist 2012;17(11):1475-1481 [FREE Full text] [doi: 10.1634/theoncologist.2012-0223] [Medline: 22933591]

10. Levine AS, Balk JL. Yoga and quality-of-life improvement in patients with breast cancer: a literature review. Int J Yoga Therap 2012(22):95-99. [Medline: 23070679]

11. Shneerson C, Taskila T, Gale N, Greenfield S, Chen Y. The effect of complementary and alternative medicine on the quality of life of cancer survivors: a systematic review and meta-analyses. Complement Ther Med 2013 Aug;21(4):417-429. [doi: 10.1016/j.ctim.2013.05.003] [Medline: 23876573]

12. Arthur K, Belliard JC, Hardin SB, Knecht K, Chen C, Montgomery S. Practices, attitudes, and beliefs associated with complementary and alternative medicine (CAM) use among cancer patients. Integr Cancer Ther 2012 Sep;11(3):232-242 [FREE Full text] [doi: 10.1177/1534735411433832] [Medline: 22313741]

13. Citrin D, Bloom D, Grutsch J, Mortensen S, Lis C. Beliefs and Perceptions of Women with Newly Diagnosed Breast Cancer Who Refused Conventional Treatment in Favor of Alternative Therapies. The Oncologist 2012 2012;17(5):612.

14. D'Arena G, Laurenti L, Coscia M, Cortelezzi A, Chiarenza A, Pozzato G. Complementary and alternative medicine use in patients with chronic lymphocytic leukemia: an Italian multicentric survey. Leuk Lymphoma 2014 Apr 1;? 2014;55(4):847. [doi: 10.3109/10428194.2013.803223]

15. Hübner J, Hanf V. Commonly Used Methods of Complementary Medicine in the Treatment of Breast Cancer. Breast Care 2013;8(5):1. [doi: 10.1159/000355705]

16. Jong MC, van de Vijver Lucy, Busch M, Fritsma J, Seldenrijk R. Integration of complementary and alternative medicine in primary care: what do patients want? Patient Educ Couns 2012 Dec;89(3):417-422. [doi: 10.1016/j.pec.2012.08.013] [Medline: 23031611]

17. Frenkel M, Ben-Arye E, Cohen L. Communication in cancer care: discussing complementary and alternative medicine. Integr Cancer Ther 2010 Jun;9(2):177-185. [doi: 10.1177/1534735410363706] [Medline: 20356950]

18. Fondazione CENSIS. 2008. Tra passato e futuro Rapporto con i farmaci e Health Technology Assessment Internet URL: http://www.censis.it/censis/census_utilities/download_content?landing_page=http%3A%2F%2Fwww.censis. it%2Fcensis%2Fbrowse%2F7%3Fshadow_comunicato_stampa%3D117754&resource=117753&resource_type=Censis%3A%3AShadowDownload; [accessed 2015-02-02] [WebCite Cache ID 6XSpP7OZb]

19. Smithson J, Paterson C, Britten N, Evans M, Lewith G. Cancer patients' experiences of using complementary therapies: polarization and integration. J Health Serv Res Policy 2010 Apr;15 Suppl 2:54-61. [doi: 10.1258/jhsrp.2009.009104] [Medline: 20194431]

20. Kim DY, Lee HR, Nam EM. Assessing cancer treatment related information online: unintended retrieval of complementary and alternative medicine web sites. Eur J Cancer Care (Engl) 2009 Jan;18(1):64-68. [doi: 10.1111/j.1365-2354.2008.00944.x] [Medline: 18771532]

21. Lee SY, Hawkins R. Why do patients seek an alternative channel? The effects of unmet needs on patients' health-related Internet use. J Health Commun 2010 Mar;15(2):152-166. [doi: 10.1080/10810730903528033] [Medline: 20390984]

22. Bjerså K, Stener VE, Fagevik OM. Knowledge about complementary, alternative and integrative medicine (CAM) among registered health care providers in Swedish surgical care: a national survey among university hospitals. BMC Complement Altern Med 2012;12:42 [FREE Full text] [doi: 10.1186/1472-6882-12-42] [Medline: 22498305]

23. Lindeman M. Biases in intuitive reasoning and belief in complementary and alternative medicine. Psychol Health 2011 Mar;26(3):371-382. [doi: 10.1080/08870440903440707] [Medline: 20419560]

24. Goldrosen M, Straus S. Complementary and alternative medicine: assessing the evidence for immunological benefits. Nat Rev Immunol 2004;4(11):921. [doi: 10.1038/nri1486]

25. Ernst E. Complementary and alternative medicine (CAM) and cancer: the kind face of complementary medicine. Int J Surg 2009 Dec;7(6):499-500 [FREE Full text] [doi: 10.1016/j.ijsu.2009.08.005] [Medline: 19716449]

26. Ahluwalia S, Murray E, Stevenson F, Kerr C, Burns J. 'A heartbeat moment': qualitative study of GP views of patients bringing health information from the internet to a consultation. Br J Gen Pract 2010 Feb;60(571):88-94 [FREE Full text] [doi: 10.3399/bjgp10X483120] [Medline: 20132702]

XSL•FO

RenderX

27. Cassileth B. Alternative Therapies: Learning from Patients Who Choose Them. The Oncologist 2012;17(5):591. [doi: 10.1634/theoncologist.2012-0149]

28. Bessi A, Coletto M, Davidescu G, Scala A, Caldarelli G, Quattrociocchi W. ArXiv Prepr. 2014 Aug 7. Science vs Conspiracy: collective narratives in the age of (mis)information URL: http://arxiv.org/pdf/1408.1667.pdf [accessed 2016-05-04] [WebCite Cache ID 6hFt7uo4q]

29. Goldacre B. Media misinformation and health behaviours. Lancet Oncol 2009 Sep;10(9):848. [doi: 10.1016/S1470-2045(09)70252-9] [Medline: 19717089]

30. Schulz PJ, Nakamoto K. Patient behavior and the benefits of artificial intelligence: the perils of "dangerous" literacy and illusory patient empowerment. Patient Educ Couns 2013 Aug;92(2):223-228. [doi: 10.1016/j.pec.2013.05.002] [Medline: 23743214]

31. Mocanu D, Rossi L, Zhang Q, Karsai M, Quattrociocchi W. ArXiv14033344 Phys. 2014 Mar 13. Collective attention in the age of (mis)information URL: http://arxiv.org/abs/1403.3344 [accessed 2016-05-04] [WebCite Cache ID 6hFtAvXUO]

32. Quattrociocchi W, Caldarelli G, Scala A. ArXiv13074292 Phys. 2013 Jul 16. Influence of media on collective debates URL: http://arxiv.org/abs/1307.4292 [accessed 2016-05-04] [WebCite Cache ID 6hFtC9fq9]

33. Huebner J, Senf B, Micke O, Muecke R, Stoll C, Prott FJ, working group PreventionIntegrative Oncology (PRIO) of the German Cancer Society. Online information on complementary and alternative medicine for cancer patients: evidence-based recommendations. Onkologie 2013;36(5):273-278. [doi: 10.1159/000350307] [Medline: 23689222]

34. Keim-Malpass J, Albrecht TA, Steeves RH, Danhauer SC. Young women's experiences with complementary therapies during cancer described through illness blogs. West J Nurs Res 2013 Nov;35(10):1309-1324 [FREE Full text] [doi: 10.1177/0193945913492897] [Medline: 23820874]

35. Price A, Grann VR. Portrayal of complementary and alternative medicine for cancer by top online news sites. J Altern Complement Med 2012 May;18(5):487-493. [doi: 10.1089/acm.2011.0110] [Medline: 22551076]

36. Walsh MC, Trentham-Dietz A, Schroepfer TA, Reding DJ, Campbell B, Foote ML, et al. Cancer information sources used by patients to inform and influence treatment decisions. J Health Commun 2010 Jun;15(4):445-463. [doi: 10.1080/10810731003753109] [Medline: 20574881]

37. DiFonzo N, Robinson NM, Suls JM, Rini C. Rumors about cancer: content, sources, coping, transmission, and belief. J Health Commun 2012;17(9):1099-1115. [doi: 10.1080/10810730.2012.665417] [Medline: 22724591]

38. CENSIS. Fondazione CENSIS. 2012. Quale futuro per il rapporto medico paziente nella nuova sanità? Internet URL: http://www.censis.it/14?shadow_ricerca=117752 [accessed 2015-04-01] [WebCite Cache ID 6XSjrWq44]

39. Reid B, Abell A, Mathews G, Rodriguez K. MDigitalLife|WCG. 2013. The Social Oncology Project 2013 Internet URL: http://www.wcgworld.com/mdigitallife/ [accessed 2015-04-01] [WebCite Cache ID 6XSpTuIO0]

40. Richards T, Montori VM, Godlee F, Lapsley P, Paul D. Let the patient revolution begin. BMJ 2013;346:f2614. [Medline: 23674136]

41. Pudota N, Dattolo A, Baruzzo A, Ferrara F, Tasso C. Automatic keyphrase extraction and ontology mining for content-based tag recommendation. Int J Intell Syst 2010;25(12):1186. [doi: 10.1002/int.20448]

42. Eysenbach G. Infodemiology and Infoveillance: Tracking Online Health Information and Cyberbehavior for Public Health. Am J Prev Med 2011 May;40(5, Supplement 2):154-158. [doi: 10.1016/j.amepre.2011.02.006]

43. Greaves F, Ramirez-Cano D, Millett C, Darzi A, Donaldson L. Use of sentiment analysis for capturing patient experience from free-text comments posted online. J Med Internet Res 2013;15(11):e239 [FREE Full text] [doi: 10.2196/jmir.2721] [Medline: 24184993]

44. Yoon S, Elhadad N, Bakken S. A practical approach for content mining of Tweets. Am J Prev Med 2013 Jul;45(1):122-129 [FREE Full text] [doi: 10.1016/j.amepre.2013.02.025] [Medline: 23790998]

45. Truccolo I, Zanini F, Bufalino R. Un codice deontologico per informare il paziente. Bibl Oggi 2009;27(2):84.

46. University of Udine Artificial Intelligence Laboratory. Intervalue. 2010. ifMONITOR URL: http://www.researchvalue.net/repository/result.html?id=272 [accessed 2015-03-25] [WebCite Cache ID 6XI2blfiN]

47. CAMbrella. CAMbrella. Communication Platform on topics related to Complementary and Alternative Medicine (CAM) in Europe Internet URL: http://www.cambrella.eu/home.php [accessed 2015-03-10] [WebCite Cache ID 6Wv8ZoVj8]

48. Office of Cancer ComplementaryAlternative Medicine. National Cancer Institute. 2012. Categories of CAM therapies-Health Information - OCCAM Internet URL: http://cam.cancer.gov/health_categories.html [accessed 2016-05-04] [WebCite Cache ID 6hFtZUH3u]

49. AIMaC, Bonucci M, Garaci A, Di LC. I trattamenti non convenzionali nel malato oncologico Internet. 2nd ed. Roma: AIMaC; 2011. URL: http://www.aimac.it/libretti-tumore/trattamenti-non-convenzionali [accessed 2015-04-01] [WebCite Cache ID 6XSq7Fl5D]

50. Broccard N, Durrer A, Frei M. Lega svizzera contro il cancro. Alternativi? Complementari??: rischi e benefici di metodi non provati in oncologia?: una guida della lega contro il cancro per pazienti e cittadini Internet. 2nd ed. Berna: Lega svizzera contro il cancro; 2008. URL: http://assets.krebsliga.ch/downloads/3951.pdf [accessed 2015-04-01] [WebCite Cache ID 6XSqAUjkc]

XSL•FO

RenderX

51. Associazione Europea Medicine Tradizionali. Naturopatia Europa. Enciclopedia di naturopatia e Discipline Bionaturali URL: http://www.naturopatiaeuropea.it/cms/Enciclopedia/Enciclopedia_Discipline_bionaturali.html; [accessed 2014-05-13] [WebCite Cache ID 6Wv8YG3Nm]

52. Wikipedia. 2013. Categoria:Medicina alternativa Internet URL: http://it.wikipedia.org/w/index.php?title=Categoria:Medicina_alternativa&oldid=57469744; [accessed 2015-03-11] [WebCite Cache ID 6XSiOnMRi]

53. Passalacqua R, Campione F, Caminiti C, Salvagni S, Barilli A, Bella M, et al. Patients' opinions, feelings, and attitudes after a campaign to promote the Di Bella therapy. Lancet 1999 Apr 17;353(9161):1310-1314. [doi: 10.1016/S0140-6736(98)10253-2] [Medline: 10218529]

54. Norsa A, Martino V. Somatostatin, retinoids, melatonin, vitamin D, bromocriptine, and cyclophosphamide in chemotherapy-pretreated patients with advanced lung adenocarcinoma and low performance status. Cancer Biother Radiopharm 2007 Feb;22(1):50-55. [doi: 10.1089/cbr.2006.365] [Medline: 17627413]

55. Todisco M, Casaccia P, Rossi N. Cyclophosphamide plus somatostatin, bromocriptin, retinoids, melatonin and ACTH in the treatment of low-grade non-Hodgkin's lymphomas at advanced stage: results of a phase II trial. Cancer Biother Radiopharm 2001 Apr;16(2):171-177. [doi: 10.1089/108497801300189263] [Medline: 11385964]

56. Report N. Risultati della sperimentazione del Multitrattamento Di Bella (MDB) Internet. Roma: Istituto Superiore di Sanità; 1999. /12 URL: http://www.iss.it/binary/publ/publi/9912.1109937886.pdf [accessed 2015-04-01] [WebCite Cache ID 6XSqMTKJC]

57. Società Italiana di Farmacologia. Position Paper della Società Italiana di Farmacologia sulla problematica Di Bella Internet. Milano: Società Italiana di Farmacologia; 2013. URL: http://www.agenziafarmaco.gov.it/sites/default/files/Position%20Paper%20SIF%20Problematica%20Di%20Bella.pdf; [WebCite Cache ID 6XSqP8AfR]

58. Glaeser E, Sunstein C. Does More Speech Correct Falsehoods? J Leg Stud 2014;43(1):93.

59. Yun Y, Lee M, Park S, Kim Y, Lee W, Lee K. Effect of complementary and alternative medicine on the survival and health-related quality of life among terminally ill cancer patients: a prospective cohort study. Ann Oncol 2013;24(2):494. [doi: 10.1093/annonc/mds469]

60. Del Vicario M, Bessi A, Zollo F, Petroni F, Scala A, Caldarelli G, et al. The spreading of misinformation online. Proc Natl Acad Sci U S A 2016 Jan 19;113(3):554-559 [FREE Full text] [doi: 10.1073/pnas.1517441113] [Medline: 26729863]

61. Mercadante S, Vitrano V. Palliative care in Italy: problem areas emerging from the literature. Minerva Anestesiol 2010 Dec;76(12):1060-1071 [FREE Full text] [Medline: 20634788]

62. Consiglio Nazionale FNOMCeO. FNOMCeO. 2009. Linee guida per la formazione nelle medicine e pratiche non convenzionali riservate ai medici chirurghi e odontoiatri della FNOMCeO Internet URL: http://www.omceomb.it/public/upload/Normativa/Linee_guida_FNOMCeO_formazione_nelle_medicine_e_pratiche_non_convenzionali.pdf; [WebCite Cache ID 6XSiwXiuy]

63. AIOM - Associazione Italiana di Oncologia Medica, CCM. I numeri del cancro in Italia. Brescia: Intermedia; 2015:2014.

64. Kelly B, Hornik R, Romantan A, Schwartz JS, Armstrong K, DeMichele A, et al. Cancer information scanning and seeking in the general population. J Health Commun 2010 Oct;15(7):734-753 [FREE Full text] [doi: 10.1080/10810730.2010.514029] [Medline: 21104503]

65. Siliquini R, Ceruti M, Lovato E, Bert F, Bruno S, De VE, et al. Surfing the internet for health information: an italian survey on use and population choices. BMC Med Inform Decis Mak 2011;11:21 [FREE Full text] [doi: 10.1186/1472-6947-11-21] [Medline: 21470435]

66. Jamieson S. Likert scales: how to (ab)use them. Med Educ 2004;38(12):1218. [doi: 10.1111/j.1365-2929.2004.02012.x]

Abbreviations

CAM: Complementary and alternative medicines
CENSIS: Centro studi investimenti sociali (center for social investment studies)
CRO: Centro di riferimento oncologico Aviano, national cancer institute
MSS: Mean sentiment score
NIH: National institute of health
OP: opinions
SPL: scientific and patients library of the Centro di Riferimento Oncologico Aviano
WD: Web documents
WS: Web sources

XSL·FO

RenderX

Original Paper

Google Flu Trends Spatial Variability Validated Against Emergency Department Influenza-Related Visits

Joseph Jeffrey Klembczyk[1], MPH, MD; Mehdi Jalalpour[2], PhD; Scott Levin[3], PhD; Raynard E Washington[4], PhD; Jesse M Pines[5], MD, MBA, MSCE; Richard E Rothman[3], MD, PhD; Andrea Freyer Dugas[3], MD, PhD

[1]Johns Hopkins University, School of Medicine, Hampstead, NC, United States

[2]Cleveland State University, Department of Civil and Environmental Engineering, Cleveland, OH, United States

[3]Johns Hopkins University, Department of Emergency Medicine, Baltimore, MD, United States

[4]Agency for Healthcare Research and Quality, Rockville, MD, United States

[5]Departments of Emergency Medicine and Health Policy, George Washington University, Washington, DC, United States

Corresponding Author:
Joseph Jeffrey Klembczyk, MPH, MD
Johns Hopkins University
School of Medicine
128 S Belvedere Dr
Hampstead, NC, 28443
United States
Phone: 1 518 573 2045
Fax: 1 910 937 1802
Email: jjklem@gmail.com

Abstract

Background: Influenza is a deadly and costly public health problem. Variations in its seasonal patterns cause dangerous surges in emergency department (ED) patient volume. Google Flu Trends (GFT) can provide faster influenza surveillance information than traditional CDC methods, potentially leading to improved public health preparedness. GFT has been found to correlate well with reported influenza and to improve influenza prediction models. However, previous validation studies have focused on isolated clinical locations.

Objective: The purpose of the study was to measure GFT surveillance effectiveness by correlating GFT with influenza-related ED visits in 19 US cities across seven influenza seasons, and to explore which city characteristics lead to better or worse GFT effectiveness.

Methods: Using Healthcare Cost and Utilization Project data, we collected weekly counts of ED visits for all patients with diagnosis (International Statistical Classification of Diseases 9) codes for influenza-related visits from 2005-2011 in 19 different US cities. We measured the correlation between weekly volume of GFT searches and influenza-related ED visits (ie, GFT ED surveillance effectiveness) per city. We evaluated the relationship between 15 publically available city indicators (11 sociodemographic, two health care utilization, and two climate) and GFT surveillance effectiveness using univariate linear regression.

Results: Correlation between city-level GFT and influenza-related ED visits had a median of .84, ranging from .67 to .93 across 19 cities. Temporal variability was observed, with median correlation ranging from .78 in 2009 to .94 in 2005. City indicators significantly associated ($P<.10$) with improved GFT surveillance include higher proportion of female population, higher proportion with Medicare coverage, higher ED visits per capita, and lower socioeconomic status.

Conclusions: GFT is strongly correlated with ED influenza-related visits at the city level, but unexplained variation over geographic location and time limits its utility as standalone surveillance. GFT is likely most useful as an early signal used in conjunction with other more comprehensive surveillance techniques. City indicators associated with improved GFT surveillance provide some insight into the variability of GFT effectiveness. For example, populations with lower socioeconomic status may have a greater tendency to initially turn to the Internet for health questions, thus leading to increased GFT effectiveness. GFT has the potential to provide valuable information to ED providers for patient care and to administrators for ED surge preparedness.

XSL•FO
RenderX

KEYWORDS

influenza; surveillance; emergency department; google flu trends; infoveillance

Introduction

Background

Influenza accounts for up to 294,000 hospitalizations and 30,000 deaths per year in the United States and costs an estimated US $12 billion annually[1-3]. Seasonal influenza patterns result in sudden increases in emergency department (ED) volume, further straining an already stressed health care safety net [4-8]. Increased influenza patient volume exacerbates ED crowding, which is linked to delays in critical treatments and increased morbidity and mortality [9-12]. Beyond seasonal influenza, the potential for a pandemic influenza outbreak is a well-recognized and serious threat to the US health care infrastructure [5,8]. Therefore, accurate and timely influenza surveillance is critical for diagnosis and treatment, as well as public health and hospital preparedness.

The Centers for Disease Control and Prevention (CDC) publicly releases weekly influenza surveillance information aggregated from diagnostic laboratories, reports from outpatient providers, and mortality and hospitalization data [4]. Although widely relied upon, the CDC surveillance information is released with a 1-2 week delay [9]. In order to provide a more timely estimate of influenza activity, Google developed Google Flu Trends (GFT), an algorithm assessing billions of Internet search queries from Google users at various geographic levels. GFT was trained with CDC regional data to estimate the proportion of outpatient visits that were related to influenza-like illness (ILI) [13]. GFT time series data can be obtained down to the city level for 122 large metropolitan areas in the United States. Although the exact algorithm calculating these estimates is proprietary, this geographically focused, publicly available data is a potential source for timely surveillance information [13].

Prior Work

Since the original validation of GFT in 2008, numerous independent evaluations have shown variable results [13]. Many studies of GFT have shown close correlation between GFT and either ILI or confirmed influenza cases in broad geographic areas and individual cities [14-16]. GFT has also been successfully included in numerous influenza forecasting models,

at both the local and the national level [17-20]. Others have identified challenges for GFT estimates. Specifically, the H1N1 pandemic in 2009 was predicted late and underestimated by GFT. This was attributed to its unusual timing and altered Internet search habits following increased media coverage of the pandemic [21,22]. Consequently, GFT's algorithm was updated to include more direct influenza-related terms rather than complications of the disease [21,22]. Even with the updated algorithm, GFT underestimated the moderately severe 2012-2013 influenza season [21]. The GFT algorithm was subsequently updated twice more in 2013 and 2014. The value of GFT and the settings in which it is most effective are not well understood.

Objective

Although there have been promising single center validations, expansion to broader geographic locations is required to fully evaluate the potential role for this alternate or complementary source of influenza surveillance. This study is the first to examine the effectiveness of GFT simultaneously in several geographically distinct regions throughout the United States. Additionally, we explored the correlation of several sociodemographic factors with GFT effectiveness. This was completed to determine the factors associated with GFT effectiveness and increase our understanding of the tool. We hypothesize that GFT will be validated as a geographically robust early predictive signal for ED influenza.

Methods

Study Population and Setting

This study, in collaboration with the Agency for Healthcare Research and Quality (AHRQ), used data from the (HCUP) State Emergency Department Databases (SEDD) to estimate influenza-related ED visits in 19 US cities from 2005-2011. The SEDD are a set of databases that include nearly all ED visits from non-rehabilitation community hospitals in participating states [23]. The 19 cities were selected based on availability of both HCUP and GFT city-level data. The cities evaluated are listed in Figure 1.

Figure 1. Correlation coefficients between Google Flu Trends and influenza-related emergency department visits.

City	2005	2006	2007	2008	2009	2010	2011	Overall
St. Paul, MN	0.964	0.849	0.844	0.809	0.621	0.942	0.935	0.821
St. Louis, MO	0.982	0.928	0.87	0.905	0.898	0.985	0.926	0.912
Salt Lake City, UT	0.977	0.943	0.828	0.961	0.78	0.95	0.708	0.825
Omaha, NE	0.875	0.839	0.903	0.931	0.769	0.96	0.862	0.835
Newark, NJ	0.721	0.845	0.752	0.816	0.378	0.826	0.67	0.672
Nashville, TN	0.953	0.808	0.936	0.953	0.876	0.868	0.961	0.875
Milwaukee, WI	0.939	0.818	0.887	0.828	0.741	0.931	0.872	0.828
Memphis, TN	0.956	0.843	0.842	0.83	0.847	0.855	0.791	0.847
Madison, WI	0.854	0.731	0.778	0.908	0.767	0.927	0.807	0.79
Lincoln, NE	0.895	0.757	0.79	0.907	0.717	0.846	0.942	0.808
Knoxville, TN	0.871	0.884	0.926	0.959	0.936	0.964	0.975	0.925
Kansas City, MO	0.934	0.798	0.872	0.957	0.781	0.937	0.929	0.844
Indianapolis, IN	0.939	0.871	0.855	0.957	0.866	0.955	0.94	0.923
Honolulu, HI	0.857	0.58	0.716	0.796	0.92	0.506	0.171	0.798
Greenville, SC	0.868	0.792	0.814	0.922	0.871	0.933	0.953	0.893
Des Moines, IA	0.971	0.898	0.863	0.865	0.762	0.975	0.891	0.877
Columbia, SC	0.937	0.855	0.782	0.863	0.905	0.882	0.968	0.854
Boston, MA	0.864	0.873	0.922	0.949	0.761	0.889	0.856	0.852
Atlanta, GA	0.94	0.927	0.932	0.977	0.847	0.811	0.899	0.761
Median	0.937	0.845	0.855	0.908	0.781	0.931	0.899	0.844
First Quartile	0.870	0.803	0.802	0.847	0.762	0.862	0.832	0.815
Third Quartile	0.955	0.879	0.895	0.955	0.874	0.953	0.941	0.876
Year	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**Overall**

Corr: 0.9, 0.8, 0.7

Data Collection

We obtained the weekly number of ED visits for influenza-related illnesses among selected cities from the HCUP databases for January 1, 2005, through December 31, 2011 [23]. This contained all ED visits to community hospitals located within the designated city area: both visits that resulted in a treatment and discharge, as well as visits that resulted in hospital admission. We defined influenza-related illness using International Statistical Classification of Diseases and Related Health Problems (ICD-9-M) codes representing diagnoses related to pneumonia or influenza (480-487, 488.1), as described by Rubison et al [24]. The addition of select pneumonia diagnoses has been validated for accurately characterizing influenza [24]. The date of the ED visit was used to create weekly totals of ED encounters for influenza-related visits for each city. Because this de-identified data was collected for another purpose, this research was exempt from the Institutional Review Board.

City-level GFT data were downloaded from the Google Flu Trends website in June 2014 for each of the 19 cities and corresponded to the 2009 update of GFT. Output consists of a local weekly parameter estimating the proportion of outpatient visits for ILI [25].

A total of 15 city indicators hypothesized to explain GFT efficacy was collected. They comprised 11 sociodemographic, two health care utilization, and two climate city-based characteristics. These measurements were most often available annually or occasionally through less-frequent surveys. The most appropriate available discrete measurement or average over our study period of the indicator was used for analysis.

The 15 sociodemographic characteristics collected for each city from the US Census Bureau (2010) included the following:

1. Population density.

2. Proportion of the population female.

3. Proportion of the population 18 years of age.

4. Proportion of the population ≥65 years of age.

5. Proportion of the population Caucasian.

6. Proportion of the population African/American.

7. Proportion of the population Hispanic/Latino [26].

8. Proportion of the population uninsured, which was collected from the 2008 Small Area Health Insurance Estimates project [27].

9. Proportion of the population with Medicare in 2008, which was collected from the Centers for Medicare and Medicaid [28].

10. Availability of Internet services (relevant to Google searches) for each city, measured by the number of Internet connections per household, was collected from the Federal Communications Commission (FCC). However, the data were binned into groups of 200 such that only categories of 0-200, 200-400, etc, per 1000 households in a given county were available [29]. Thus, we used the midpoint of each bin to provide a household-weighted average among counties for each city.

11. A collective measure of socioeconomic status (SES) was created by combining four separate indicators: household median income (US Census Bureau 2010), proportion with high school degree (American Community Survey 2007-2009), proportion

with college degree (American Community Survey 2007-2009), and proportion employed (Bureau of Labor Statistics 2008, collected by county and population-weighted) [26,30,31]. These individual indicators were highly correlated and thus considered proxies for socioeconomic status. The four indicators were normalized along the 19 observations to produce the SES variable with a mean of zero (SD 3.15).

12. Medicaid-reimbursed hospital inpatient days per 1000 person-years, collected from the American Hospital Association [32].

13. Total ED visits per person-year, retrieved for 2011 from HCUP [23]. Because no significant time variation in total ED visits was observed, the temporal average was used for each city. These two health care utilization measures were available only by county, so population-weighted averages of the counties composing each city were calculated.

14. City climate conditions were included. Air pollution (particulate matter 2.5) was collected from the CDC for 2008 at the county level and was also population-weighted [33].

15. Seasonality of climate was estimated using daily historical temperature readings for each city collected from Weather Underground [34]. Average monthly temperatures along the entire time series were calculated, and the standard deviation of these averages was taken as a measure of seasonality of temperature as described by Legates and Willmott [35].

Google Flu Trends Effectiveness

Pearson correlation coefficients between GFT and ED visits for pneumonia and influenza for each city were calculated both for individual seasons and the entire time series. Each season included data from August 1 of the prior year to July 31 of the stated year with the exception of 2005, which began at January 1 and ended July 31 due to data availability. For example, the

2006 season includes data from August 1, 2005, to July 31, 2006.

We used two separate methods to identify potential outliers with respect to GFT effectiveness. First, we used the traditional box and whisker method in which cities with a correlation coefficient the distance of 1.5 times the interquartile range (IQR) outside of the IQR were considered outliers. We also applied the median absolute deviation method of outlier identification [36,37].

City Indicators

Univariate linear regression was performed along the 19 cities using each of the 15 sociodemographic variables as independent variables and the correlations between GFT and ED visits for pneumonia and influenza as the dependent variable. Trend lines were displayed only for those sociodemographic factors for which regression yielded P .10.

Results

Google Flu Trends Effectiveness

Overall, GFT is highly correlated with ED visits for pneumonia and influenza, with a median correlation of .844 (range .672-.925) across the 19 cities included in this analysis (Figure 1). However, there is temporal variability (Figure 2), with yearly median correlations ranging from .781 during the 2009 H1N1 pandemic to .937 in 2005. There is additional geographic variability, as shown in Figure 3, with a trend of higher correlations between GFT and ED visits for influenza-related visits in the midwest and southeast regions including Des Moines, IA; St. Louis, MO; Indianapolis, IN; Nashville, TN; Knoxville, TN; and Greenville, SC. Figure 4 displays a time series comparison of GFT and weekly influenza-related visits for the three cities with the lowest, median, and highest correlation.

Figure 2. Correlation between Google Flu Trends and influenza-related emergency department visits for individual cities, by year (outliers are marked by red +, including Honolulu, HI [Hnl] and Newark, NJ [Nrk]).
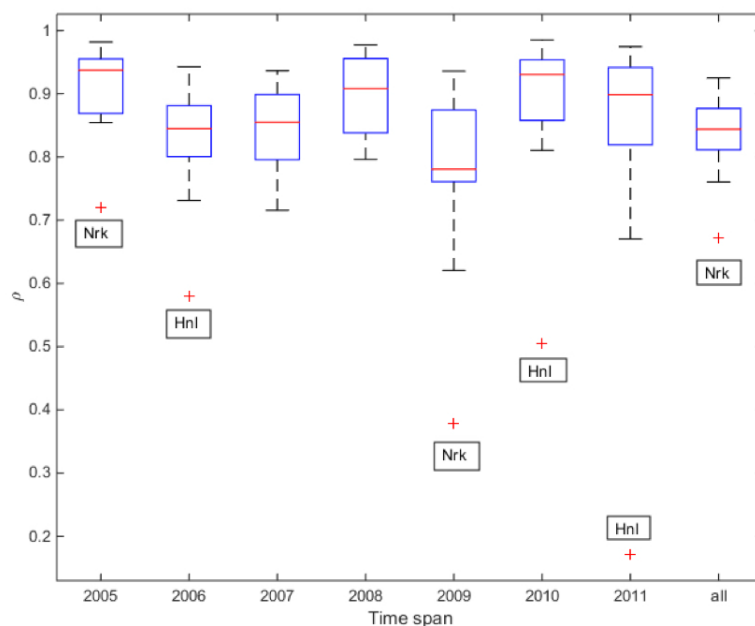
Figure 3. Correlation coefficients between Google Flu Trends and influenza-related emergency department visits for individual cities over the total time series (2005-2011). Correlations range from .672 (yellow) to .925 (red).

XSL•FO
RenderX
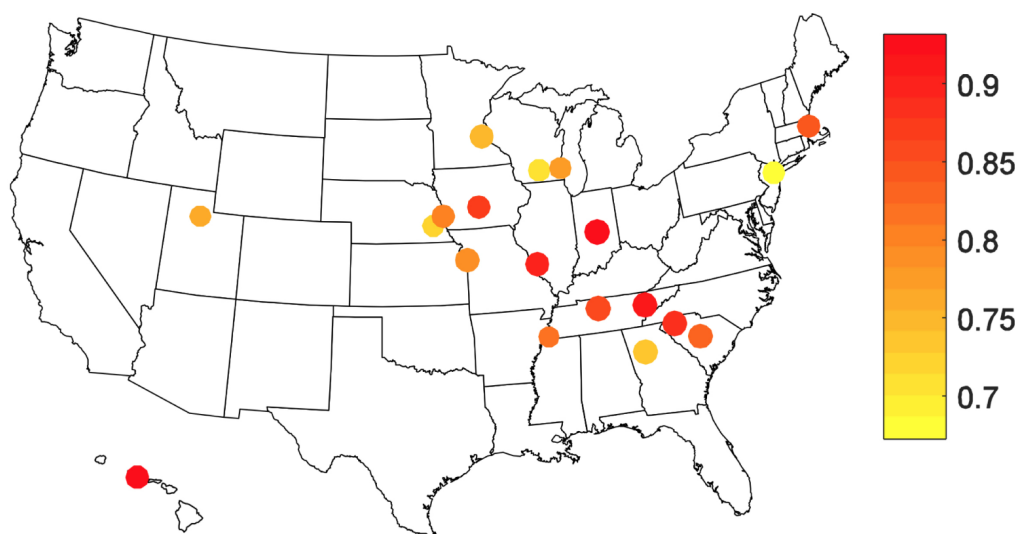
Figure 4. Time series comparing Google Flu Trends and influenza-related emergency department visits for individual cities over the total time series (2005-2011) demonstrating the lowest (Newark, NJ P=.672), median (Kansas City, MO P=.844), and highest (Knoxville, TN P=.925) correlation coefficients.
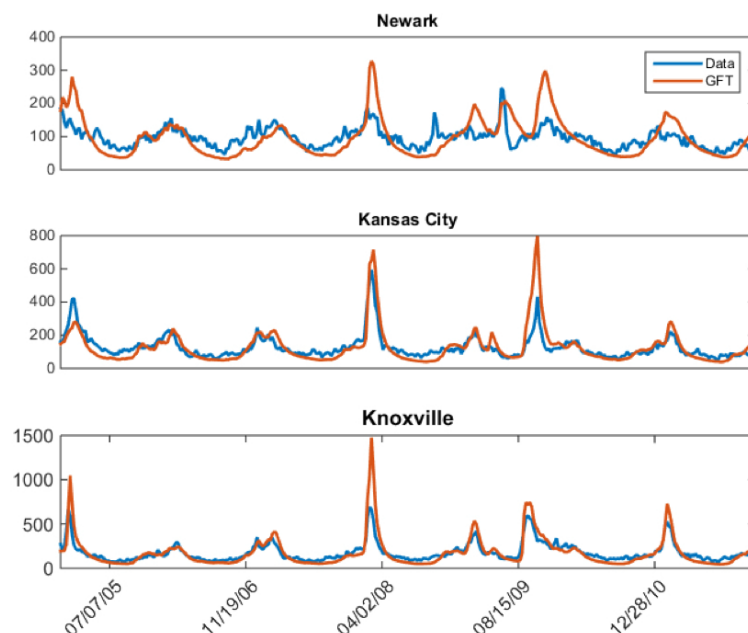
City Indicators

Newark, NJ, was found to be an outlier with respect to GFT effectiveness. This was based on consensus of the two independent outlier-identification techniques as well as expert opinion. We believe it carried undue influence in our analysis of city-based indicators and therefore removed Newark from these calculations. Honolulu appears outside the IQR for the distribution of cities in Figure 2 in three different years, but it was not quantitatively identified as an outlier over the whole time series and thus was included in the analysis.

Fifteen sociodemographic indicators, collected for each of the 19 cities, were evaluated for their potential association with the correlation between GFT and ED visits for pneumonia and influenza (Figure 5). Of the indicators evaluated, Internet availability and socioeconomic status were negatively correlated with GFT effectiveness (decrease in these variables was associated with an increase in GFT effectiveness). Proportion of the population that is female, proportion with Medicare insurance, and number of ED visits per person were positively correlated with GFT effectiveness.

Figure 5. Correlation between Google Flu Trends and emergency department visits for pneumonia and influenza for individual cities over the total time series (2005-2011) plotted against 15 different city-level indicators. Trend lines are plotted for variables with a *P* value of less than .10. Univariate regression coefficients and *P* values are displayed for each indicator. The outlier city (Newark) was not included in the analysis but is still displayed in gray.
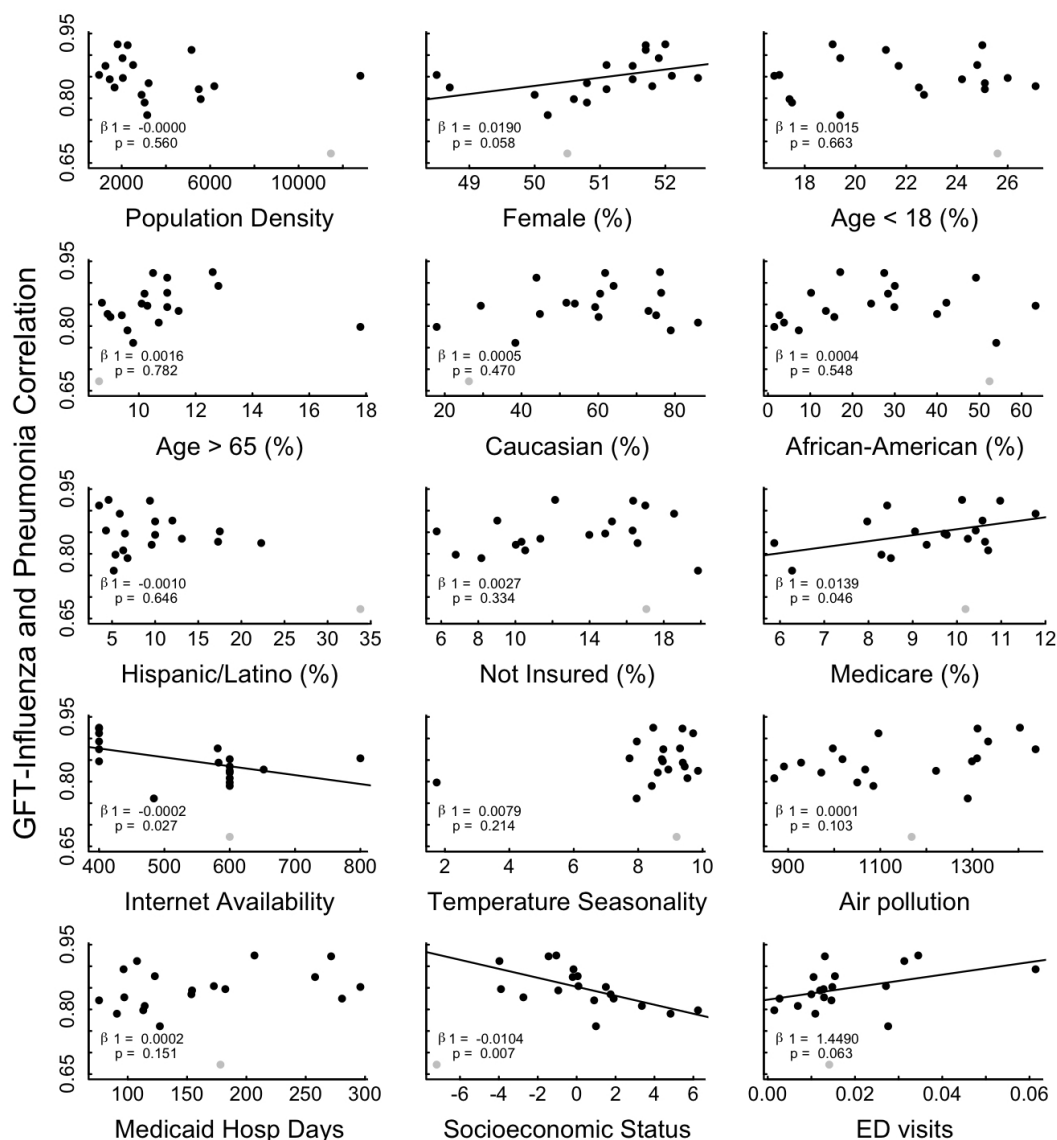
Discussion

Principal Results and Prior Work

Although GFT is a promising new source of real-time influenza surveillance, there is conflicting evidence regarding its accuracy. Previous studies have validated GFT at the national level or in a specific local setting, but this is the first to evaluate GFT across multiple cities simultaneously with local clinical outcomes [13-16,21-22]. We sought to more fully understand the geographical and temporal correlations between GFT and influenza-related ED visits by evaluating 19 different US cities over 7 influenza seasons.

Overall, we found GFT to be a valuable tool that provides useful surveillance in a variety of settings. However, there remains some geographic and temporal variability. Cities in the Southeast and Midwest appeared to have stronger correlations between GFT and influenza-related ED visits compared to cities in other

regions. Similar to our results, temporal variability in GFT effectiveness has been observed in past studies [15-20,22]. This may be due to a combination of outbreak timing, outbreak severity, media coverage, public health awareness, and other unpredictable sources of variability. GFT has been updated in the past in an attempt to reduce some of these problems [21]. Characterization and minimization of this temporal variability is critical when incorporating GFT into influenza surveillance systems.

We further explored the geographic findings by evaluating characteristics of individual cities that may impact the relationship between GFT and influenza-related ED visits. The only basic city demographic variable that correlated with effective GFT was proportion of the population that is female. Per-capita health care use tends to be higher among females, which may explain this trend [38,39]. More notable was that several factors including age and ethnicity did not correlate with GFT. We hypothesized that proportionally older populations

(age 65 years) may be less likely to access health information on the Internet; however, the proportion of those populations in a city did not impact GFT effectiveness. Additionally, there was no change in GFT surveillance effectiveness in cities with a large Hispanic or Latino population despite the hypothesis that primary language differences may limit search queries counted by the GFT algorithm, which uses English search terms only.

Other indicators are more difficult to interpret. Internet connections per household was associated with decreased GFT effectiveness, while we hypothesized that greater connectivity would lead to more predictive GFT. Internet availability was only available as data binned into 5 levels from the FCC. The granularity in measurement of this variable may have limited its utility in accurately distinguishing differences between the cities. Furthermore, our hypothesis would be best tested by a measure of Internet use, rather than availability, but a consistent indicator of use was not readily available.

Lower SES was associated with more effective GFT. This may be because lower SES populations may disproportionately use the ED for non-urgent conditions (eg, ILI) due to limited access to other community health services such as primary care [40-44]. This SES effect is likely more than a reflection of the health insurance status of the populations, as the correlation with proportion uninsured was insignificant. Further, lower SES populations may be more likely to consult the Internet for health care questions, resulting in more accurate GFT predictions.

In evaluating the correlation between health utilization with GFT effectiveness, both the proportion of the population insured by Medicare and the per capita number of ED visits had a positive correlation with GFT effectiveness. Given that we are evaluating GFT effectiveness through correlation with ED visits, it is expected that cities more dependent upon ED care would have stronger correlation between a marker of influenza and ED visits for potential influenza. Therefore, we would expect GFT to be most useful as an ED and hospital surveillance tool in populations with lower SES and higher ED utilization.

Markers of local climate, such as air pollution or seasonality of climate, did not correlate with the effectiveness of GFT. Several influenza forecast models have included temperature to predict severity of influenza [45,46]. Cities with increased variation in temperature by season may have more severe and predictable influenza. However, the insignificance of the climate variable, and the determination that warmer cities in the southeast United States had increased GFT effectiveness, fail to support this hypothesis.

Similarly, we hypothesized that cities with increased air pollution might have poorer baseline lung function and thus more severe influenza pathology. This would cause heightened influenza awareness and diagnostic rates, leading to improved GFT effectiveness. However, this effect was also not supported by our data. Our analysis suggests that GFT effectiveness may not be driven by severity of disease.

Our results support the conclusion that traditional surveillance models can benefit from the addition of Internet search query data. However, temporal and geographic variability exists, which should be considered when generalizing results from a single influenza season or single hospital or region. This study specifically demonstrates the magnitude of variability that may be expected across different cities in the United States. Further, our results suggest that a population-based measure of SES may be useful to understand and modulate confidence in GFT effectiveness. Regardless, before incorporating GFT or other Internet query-based data into local public health surveillance systems, it is important to account for GFT performance in that specific location.

Limitations

Limitations of our study include a small sample size of 19 cities, which may have hindered our ability to detect trends in city characteristics. The sample size also constrained us from carrying out multivariate regression analyses. Additionally, historical GFT data were available only in weekly intervals, limiting the temporal resolution of our analyses. As previously mentioned, Internet access and usage was difficult to quantify. Health care access and utilization was also difficult to capture at the local level, and more available variables in this category may have yielded further insight. Additionally, Newark was excluded as an outlier from the sociodemographic factors analysis. While we justified the decision to remove Newark, it did affect the significance of some trends: proportion Hispanic/Latino became insignificant, while SES became significant. The sensitivity of our results is a function of both the small sample size as well as the extreme values for Newark in GFT effectiveness and some of the city-level indicators. Next, while we validated GFT's correlation with influenza-related ED visits, GFT is more broadly designed to correlate with outpatient ILI visits. Therefore, our inferences of the factors driving GFT effectiveness may not be generalizable to GFT as used in settings outside of the emergency department. Moreover, the study used ED visits data up to 2011 and the corresponding 2009 GFT model of the era, which limits generalizing the conclusions about GFT to recent trends. Finally, while GFT access is currently limited by Google to only research institutions, our results are still relevant to future iterations of GFT and other Internet search query-based surveillance tools.

Conclusions

As a whole, our results indicate that GFT is a sensitive surveillance tool that can add value to our current surveillance systems. Because of its spatio-temporal variability in effectiveness, GFT is likely most useful as an additional, early signal to influenza prediction models, rather than as a stand-alone approach. Furthermore, our results help explain where GFT may be most effective, specifically in higher percent female populations with lower socioeconomic status and high ED use. This can help inform the most useful settings for further GFT study and implementation. Effective, real-time influenza surveillance is useful both for emergency medicine providers on a patient-to-patient basis and for ED crowding preparedness. Characterizing geographic effectiveness and variability of GFT and Internet search query data is crucial for the continued progress of influenza surveillance.

Conflicts of Interest

None declared.

References

1. Centers for Disease Control and Prevention. Seasonal influenza. 2011. URL: http://www.cdc.gov/flu/about/qa/disease.htm [accessed 2016-01-28] [WebCite Cache ID 6erU01w1f]
2. Thompson WW, Shay DK, Weintraub E, Brammer L, Bridges CB, Cox NJ, et al. Influenza-associated hospitalizations in the United States. JAMA 2004 Sep 15;292(11):1333-1340. [doi: 10.1001/jama.292.11.1333] [Medline: 15367555]
3. Thompson WW, Shay DK, Weintraub E, Brammer L, Cox N, Anderson LJ, et al. Mortality associated with influenza and respiratory syncytial virus in the United States. JAMA 2003 Jan 8;289(2):179-186. [Medline: 12517228]
4. McDonnell WM, Nelson DS, Schunk JE. Should we fear "flu fear" itself? Effects of H1N1 influenza fear on ED use. Am J Emerg Med 2012 Feb;30(2):275-282. [doi: 10.1016/j.ajem.2010.11.027] [Medline: 21208765]
5. Institute of Medicine Committee on the Future of Emergency Care in the U.S. Health System. The future of emergency care in the United States health system. Ann Emerg Med 2006 Aug;48(2):115-120. [doi: 10.1016/j.annemergmed.2006.06.015] [Medline: 16857459]
6. Glaser CA, Gilliam S, Thompson WW, Dassey DE, Waterman SH, Saruwatari M, et al. Medical care capacity for influenza outbreaks, Los Angeles. Emerg Infect Dis 2002 Jun;8(6):569-574 [FREE Full text] [doi: 10.3201/eid0806.010370] [Medline: 12023911]
7. Schull MJ, Mamdani MM, Fang J. Community influenza outbreaks and emergency department ambulance diversion. Ann Emerg Med 2004 Jul;44(1):61-67. [doi: 10.1016/S0196064403013167] [Medline: 15226710]
8. Osterholm MT. Preparing for the next pandemic. N Engl J Med 2005 May 5;352(18):1839-1842. [doi: 10.1056/NEJMp058068] [Medline: 15872196]
9. Bernstein SL, Aronsky D, Duseja R, Epstein S, Handel D, Hwang U, Society for Academic Emergency Medicine, Emergency Department Crowding Task Force. The effect of emergency department crowding on clinically oriented outcomes. Acad Emerg Med 2009 Jan;16(1):1-10 [FREE Full text] [doi: 10.1111/j.1553-2712.2008.00295.x] [Medline: 19007346]
10. Pines JM, Hollander JE, Localio AR, Metlay JP. The association between emergency department crowding and hospital performance on antibiotic timing for pneumonia and percutaneous intervention for myocardial infarction. Acad Emerg Med 2006 Aug;13(8):873-878 [FREE Full text] [doi: 10.1197/j.aem.2006.03.568] [Medline: 16766743]
11. Pines JM, Pollack CV, Diercks DB, Chang AM, Shofer FS, Hollander JE. The association between emergency department crowding and adverse cardiovascular outcomes in patients with chest pain. Acad Emerg Med 2009 Jul;16(7):617-625 [FREE Full text] [doi: 10.1111/j.1553-2712.2009.00456.x] [Medline: 19549010]
12. Pines JM, Shofer FS, Isserman JA, Abbuhl SB, Mills AM. The effect of emergency department crowding on analgesia in patients with back pain in two hospitals. Acad Emerg Med 2010 Mar;17(3):276-283 [FREE Full text] [doi: 10.1111/j.1553-2712.2009.00676.x] [Medline: 20370760]
13. 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]

14. Dugas AF, Hsieh Y, Levin SR, Pines JM, Mareiniss DP, Mohareb A, et al. Google Flu Trends: correlation with emergency department influenza rates and crowding metrics. Clin Infect Dis 2012 Feb 15;54(4):463-469 [FREE Full text] [doi: 10.1093/cid/cir883] [Medline: 22230244]

15. Thompson LH, Malik MT, Gumel A, Strome T, Mahmud SM. Emergency department and 'Google flu trends' data as syndromic surveillance indicators for seasonal influenza. Epidemiol Infect 2014 Nov;142(11):2397-2405. [doi: 10.1017/S0950268813003464] [Medline: 24480399]

16. Araz OM, Bentley D, Muelleman RL. Using Google Flu Trends data in forecasting influenza-like-illness related ED visits in Omaha, Nebraska. Am J Emerg Med 2014 Sep;32(9):1016-1023. [doi: 10.1016/j.ajem.2014.05.052] [Medline: 25037278]

17. Dugas AF, Jalalpour M, Gel Y, Levin S, Torcaso F, Igusa T, et al. Influenza forecasting with Google Flu Trends. PLoS One 2013;8(2):e56176 [FREE Full text] [doi: 10.1371/journal.pone.0056176] [Medline: 23457520]

18. Pervaiz F, Pervaiz M, Abdur RN, Saif U. FluBreaks: early epidemic detection from Google flu trends. J Med Internet Res 2012;14(5):e125 [FREE Full text] [doi: 10.2196/jmir.2102] [Medline: 23037553]

19. Shaman J, Karspeck A. Forecasting seasonal outbreaks of influenza. Proc Natl Acad Sci U S A 2012 Dec 11;109(50):20425-20430 [FREE Full text] [doi: 10.1073/pnas.1208772109] [Medline: 23184969]

20. Preis T, Moat HS. Adaptive nowcasting of influenza outbreaks using Google searches. R Soc Open Sci 2014 Oct;1(2):140095 [FREE Full text] [doi: 10.1098/rsos.140095] [Medline: 26064532]

21. Olson DR, Konty KJ, Paladini M, Viboud C, Simonsen L. Reassessing Google Flu Trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales. PLoS Comput Biol 2013;9(10):e1003256 [FREE Full text] [doi: 10.1371/journal.pcbi.1003256] [Medline: 24146603]

22. Cook S, Conrad C, Fowlkes AL, Mohebbi MH. Assessing Google flu trends performance in the United States during the 2009 influenza virus A (H1N1) pandemic. PLoS One 2011;6(8):e23610 [FREE Full text] [doi: 10.1371/journal.pone.0023610] [Medline: 21886802]

23. Agency for Healthcare Research and Quality. September. 2014. Healthcare Cost and Utilization Project (HCUP) Overview of the State Emergency Department Databases URL: http://www.hcup-us.ahrq.gov/db/state/sedddist/SEDD_Introduction.jsp [accessed 2016-01-28] [WebCite Cache ID 6erUqnDGv]

24. Rubinson L, Mutter R, Viboud C, Hupert N, Uyeki T, Creanga A, et al. Impact of the fall 2009 influenza A(H1N1)pdm09 pandemic on US hospitals. Med Care 2013 Mar;51(3):259-265. [doi: 10.1097/MLR.0b013e31827da8ea] [Medline: 23295577]

25. Google Flu Trends. 2014. URL: https://www.google.org/flutrends/about/data/flu/historic/us-historic-v2.txt [accessed 2014-06-05] [WebCite Cache ID 6erVyQWjb]

26. United States Census Bureau. United States Census Bureau Quick Facts. 2010. Washington: GPO URL: http://www.census.gov/quickfacts/ [accessed 2016-06-15] [WebCite Cache ID 6iIRvTQTF]

27. United States Census Bureau; Small Area Health Insurance Estimates. 2008. Washington: GPO URL: http://www.census.gov/did/www/sahie/data/20082013/index.html [accessed 2016-01-28] [WebCite Cache ID 6erWFeNES]

28. Centers for Medicare and Medicaid. Medicare Geographic Variation. 2008. URL: http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Geographic-Variation/ [accessed 2016-06-15] [WebCite Cache ID 6iIS48vSo]

29. Federal Communications Commission. 2008. County Data on Internet Access Services URL: https://www.fcc.gov/general/form-477-county-data-internet-access-services [accessed 2016-01-28] [WebCite Cache ID 6erWq9Ncw]

30. United States Census Bureau. American Community Survey 2007-2009. 2009. URL: http://factfinder.census.gov/ [WebCite Cache ID 6erWzT3ik]

31. Bureau of Labor Statistics. 2008. Unemployment County Data URL: http://www.bls.gov/lau/ [accessed 2016-01-28] [WebCite Cache ID 6erXY8Dbh]

32. American Hospital Association. 2010. AHA Survey (via Area Health Resource File) URL: http://ahrf.hrsa.gov/download.htm [accessed 2016-01-28] [WebCite Cache ID 6erXf84tQ]

33. Centers for Disease Control. 2008. WONDER Online Database URL: http://wonder.cdc.gov/nasa-pm.html [accessed 2016-01-28] [WebCite Cache ID 6erXoCJVV]

34. Weather Underground. Historical Weather. URL: http://www.wunderground.com/weather/api [accessed 2016-01-28] [WebCite Cache ID 6erY0WjcK]

35. Legates DR, Willmott CJ. Mean seasonal and spatial variability in global surface air temperature. Theor Appl Climatol 1990;41(1-2):11-21. [doi: 10.1007/BF00866198]

36. Leys C, Ley C, Klein O, Bernard P, Licata L. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. Journal of Experimental Social Psychology 2013 Jul;49(4):764-766. [doi: 10.1016/j.jesp.2013.03.013]

37. Martinez W, Martinez A. Computational Statistics Handbook with MATLAB, Second Edition (Chapman & Hall/Crc Computer Science & Data Analysis). Boca Raton, FL: Chapman & Hall/CRC; 2007.

38. Bertakis KD, Azari R, Helms LJ, Callahan EJ, Robbins JA. Gender differences in the utilization of health care services. J Fam Pract 2000 Feb;49(2):147-152. [Medline: 10718692]

39. Green CA, Pope CR. Gender, psychosocial factors and the use of medical services: a longitudinal analysis. Soc Sci Med 1999 May;48(10):1363-1372. [Medline: 10369437]

40.   DeNavas-Walt C, Proctor B, Smith J. Income, Poverty, and Health Insurance Coverage in the United States. Washington, DC: U.S. Census Bureau; 2013. URL: https://www.census.gov/prod/2013pubs/p60-245.pdf [accessed 2016-06-15] [WebCite Cache ID 6iISt2zWB]

41.   Newton MF, Keirns CC, Cunningham R, Hayward RA, Stanley R. Uninsured adults presenting to US emergency departments: assumptions vs data. JAMA 2008 Oct 22;300(16):1914-1924. [doi: 10.1001/jama.300.16.1914] [Medline: 18940978]

42.   Asplin BR, Rhodes KV, Levy H, Lurie N, Crain AL, Carlin BP, et al. Insurance status and access to urgent ambulatory care follow-up appointments. JAMA 2005 Sep 14;294(10):1248-1254. [doi: 10.1001/jama.294.10.1248] [Medline: 16160133]

43.   Tang N, Stein J, Hsia RY, Maselli JH, Gonzales R. Trends and characteristics of US emergency department visits, 1997-2007. JAMA 2010 Aug 11;304(6):664-670 [FREE Full text] [doi: 10.1001/jama.2010.1112] [Medline: 20699458]

44.   Schuur JD, Venkatesh AK. The growing role of emergency departments in hospital admissions. N Engl J Med 2012 Aug 2;367(5):391-393. [doi: 10.1056/NEJMp1204431] [Medline: 22784039]

45.   Lowen AC, Mubareka S, Steel J, Palese P. Influenza virus transmission is dependent on relative humidity and temperature. PLoS Pathog 2007 Oct 19;3(10):1470-1476 [FREE Full text] [doi: 10.1371/journal.ppat.0030151] [Medline: 17953482]

46.   Sloan C, Moore ML, Hartert T. Impact of pollution, climate, and sociodemographic factors on spatiotemporal dynamics of seasonal respiratory viruses. Clin Transl Sci 2011 Feb;4(1):48-54 [FREE Full text] [doi: 10.1111/j.1752-8062.2010.00257.x] [Medline: 21348956]

Abbreviations

AHRQ: Agency for Healthcare Research and Quality
CDC: Centers for Disease Control and Prevention
ED: emergency department
FCC: Federal Communications Commission
GFT: Google Flu Trends
HCUP: Healthcare Cost and Utilization Project
ILI: influenza-like illness
IQR: interquartile range
SEDD: State Emergency Department Databases
SES: socioeconomic status

XSL•FO
RenderX

Original Paper

Health Information Brokers in the General Population: An Analysis of the Health Information National Trends Survey 2013-2014

Sarah L Cutrona[1,2,3], MD, MPH; Kathleen M Mazor[1,2], EdD; Amenah A Agunwamba[4], ScD, MPH; Sruthi Valluri[5,6], BS; Patrick M Wilson[4], MPH; Rajani S Sadasivam[3], PhD; Lila J Finney Rutten[4], PhD, MPH

[1]University of Massachusetts Medical School, Department of Medicine, Worcester, MA, United States

[2]Meyers Primary Care Institute, Worcester, MA, United States

[3]University of Massachusetts Medical School, Department of Quantitative Health Sciences, Worcester, MA, United States

[4]Robert D. and Patricia E. Kern Center for the Science of Healthcare Delivery, Mayo Clinic, Rochester, MN, United States

[5]University of Minnesota Medical School, Minneapolis, MN, United States

[6]University of Minnesota School of Public Health, Department of Epidemiology and Community Health, Minneapolis, MN, United States

Corresponding Author:
Sarah L Cutrona, MD, MPH
University of Massachusetts Medical School
Department of Medicine
365 Plantation St
Biotech 1 Suite 100
Worcester, MA, 01605
United States
Phone: 1 508 856 3086
Fax: 1 508 856 5024
Email: Sarah.Cutrona@umassmemorial.org

Abstract

Background: Health information exchanged between friends or family members can influence decision making, both for routine health questions and for serious health issues. A health information broker is a person to whom friends and family turn for advice or information on health-related topics. Characteristics and online behaviors of health information brokers have not previously been studied in a national population.

Objective: The objective of this study was to examine sociodemographic characteristics, health information seeking behaviors, and other online behaviors among health information brokers.

Methods: Data from the Health Information National Trends Survey (2013-2014; n=3142) were used to compare brokers with nonbrokers. Modified Poisson regression was used to examine the relationship between broker status and sociodemographics and online information seeking.

Results: Over half (54.8%) of the respondents were consulted by family or friends for advice or information on health topics (ie, they acted as health information brokers). Brokers represented 54.1% of respondents earning <$20,000 yearly and 56.5% of respondents born outside the United States. Women were more likely to be brokers (PR 1.34, 95% CI 1.23-1.47) as were those with education past high school (PR 1.42, CI 1.22-1.65). People aged ≥75 were less likely to be brokers as compared to respondents aged 35-49 (PR 0.81, CI 0.67-0.99). Brokers used the Internet more frequently for a variety of online behaviors such as seeking health information, creating and sharing online content, and downloading health information onto a mobile device; and also reported greater confidence in obtaining health information online.

Conclusions: More than 50% of adults who responded to this national survey, including those with low income and those born abroad, were providing health information or advice to friends and family. These individuals may prove to be effective targets for initiatives supporting patient engagement and disease management, and may also be well-positioned within their respective social networks to propagate health messages.

(*J Med Internet Res 2016;18(6):e123*)   doi:10.2196/jmir.5447

Introduction

Health information exchanged between friends or family members influences decision making, both for routine health questions and for serious health concerns [1-5]. In 2014, 60% of Americans reported that they obtained health information or support from friends and family for a difficult health issue [5]. While health care professionals remain the preferred source of information for many technical questions, family and friends offer factual health information and also emotional support, drawing on personal experiences, beliefs, and attitudes [3,5]. Recognizing and supporting those who provide health care advice to their peers (both online and offline) may be an effective way to disseminate health messages to broader audiences.

Various terms have been used to describe the roles played by laypeople while providing health information to family or friends. Studies have described surrogate seekers (those who self-report seeking health information on behalf of someone other than themselves) and lay information mediaries (nonprofessionals who seek information on behalf of others without necessarily being asked to do so); related concepts include "gatekeeping," "proxy information seeking," "sharing information found for others on the Web," "information-acquiring and information-sharing," or completion of an "imposed query" or "gift query" [6-11].

These terms have been used to convey slightly different meanings, but all emphasize the act of seeking and acquiring health information before passing it on. Previous literature has described use of family, friends, or other lay interpersonal contacts as sources of health information for a broad range of topics [3,12]. Interpersonal sources of health information tend to be female, in good or excellent health, living in shared household arrangements (marriage, living with others, providing care to an adult relative) and tend to be related to someone with a serious or chronic medical condition [11,13,14]. These individuals may engage in online activities requiring user-generated content (eg, email communication with health care providers, participation in online health support groups) and often work to help patients overcome information-seeking barriers [14,15]. Those comfortable with online communication may act as intermediaries for members of at-risk populations who lack the ability (due to language, literacy, cognitive challenges, or ease with technology) to search and access online information, bridging the "digital divide" that has been described among traditionally disadvantaged groups [16]. Many factors may motivate a person to seek health information on behalf of others. This activity can stem from feelings of empathy, altruism, and a desire to help; helping behavior may also be pleasurable and thus benefit the helper [10,15,17].

Health information acquired through personal experience can also be of instrinsic value to social network members, whether passed along verbally or organized and made more accessible through the Internet [1]. Previous research defines individuals with knowledge on a health-related subject (and who consider it important to share this knowledge with others) as "health information mavens" [18]. Among low-income Massachusetts respondents, mavenism has been associated with certain characteristics (being female, older, with larger social networks and with moderate consumption of general media) [18]. In one study, mavens had also spent fewer years in the United States and had lower language acculturation levels [18]. While this study provides valuable information about a selected group of individuals, these findings may not be generalizable to a broader population. Further studies are needed to better characterize those persons who act as sources of health information for friends and family.

Using data from the Health Information National Trends Survey (HINTS) [19], we identified respondents who were acting as brokers of health information. A "health information broker" is a person to whom friends and family turn for information or advice on health related topics. We sought to describe sociodemographic characteristics, health information seeking behaviors (or lack thereof), and online health information communication preferences among brokers compared with nonbrokers. In order to understand whether our findings were applicable among traditionally disadvantaged groups, we additionally studied characteristics of brokers among respondents with low incomes and among those born outside the United States. Understanding characteristics of health information brokers may have implications related to the design of health communication campaigns, including future eHealth and mHealth interventions targeting these users.

Methods

Data Collection and Response Rates

Data for these analyses were obtained from the HINTS, a national survey of the US adult population that assesses knowledge, attitudes, and behavior related to health communication and related outcomes [19]. HINTS 4, Cycle 3 is the only version of HINTS fielded thus far to specifically ask whether "family members and friends ask you for information or advice on health topics". Data for HINTS 4, Cycle 3 were collected from September 6, 2013 to December 30, 2013 (n=12,010) through mailed questionnaires. The sample design for HINTS 4, Cycle 3 was a two-stage, stratified sample, wherein addresses were selected from a comprehensive national residential file from the United States Postal Service, and individual respondents were selected for each sampled household. The final response rate was 35.19% (n=3142) [20]. Details on sampling strategies and survey design are available in the HINTS 4, Cycle 3 methodology report [21]. HINTS 4 was approved by the Westat Institutional Review Board (IRB) in an expedited review in 2010, and was deemed exempt from IRB review by the National Institutes of Health (NIH) Office of Human Subjects Research in 2011.

Measures

Sociodemographic Variables

The following sociodemographic variables were included in these analyses: sex (male, female), age (18-34, 35-49, 50-64, 65-74, and 75+ years), race or ethnicity (non-Hispanic white, non-Hispanic black, Hispanic or Latino, and non-Hispanic "other"), annual household income (< $20K, $20 to <$35K,

$35K to <$50, $50K to <$75K, $75K or more), education (less than high school or high school graduate, some college, college/Bachelor's degree), born in the United States (yes, no), speaks English (not at all or not well, well, very well), and marital status (married or living as married, not married).

Health Information Broker Status

Each respondent was asked "Do family members and friends ask you for information or advice on health topics?" Those who responded affirmatively were classified as health information brokers.

Health Information Seeking and Sources

To assess health information seeking behavior, respondents were asked "Have you ever looked for information about health or medical topics from any source?" Those who answered "yes" were asked "For whom was the information sought during the most recent search (myself, someone else, both)?" and "The most recent time you looked for information about health or medical topics, where did you go first?" Responses were coded as family or friends, health care provider, Internet, print materials, and other sources.

On a 4-point scale, ranging from "Not at all" to "A lot," respondents were asked to rate the extent to which they trusted information about health or medical topics from the following 8 sources: doctor, family or friends, news, radio, Internet, television, government health agencies, and charitable /religious organizations. The question did not address the language in which this information was delivered. Responses were dichotomized into a lot versus all other responses. For news sources, a respondent who indicated that they had 'a lot' of trust in online newspapers, print newspapers, special health magazines, medical magazines or newsletters was categorized as having 'a lot' of trust in information from news sources; for television, a respondent who indicated that they had 'a lot' of trust in local or national television was categorized as having 'a lot' of trust in information from television.

Information Seeking Experiences

All respondents were asked to rate their degree of confidence in their ability to obtain necessary health or medical information on a scale ranging from "completely confident" to "not confident at all." For this analysis, responses were dichotomized as: completely or very confident and somewhat, a little, or not at all confident.

Each respondent was asked about their information source preferences with the following question: "Imagine that you had a strong need to get information about health or medical topics. Where would you go first?" Responses were coded as printed materials, family or friends, Internet, healthcare provider, and "other."

Internet Use

Patients were questioned on their use of the Internet; those who indicated that they use the Internet were asked about their online activities including social networking site visits, sharing health information on social networking sites, writing in an online diary or blog, participating in an online support group related to health issues, and viewing health-related YouTube videos. Online activities in the last 12 months were also documented, which included seeking health or medical information on behalf of oneself or others, seeking information on smoking cessation, purchasing of medicine or vitamins online, seeking a health care provider online, downloading health information to a mobile device, tracking personal health information, and communicating online with a doctor or with personnel in the doctor's office.

Data Analyses

Analysis of the complex survey data was conducted using SAS 9.3 and the survey package for R version 3.02 [22,23]. All data were weighted to provide estimates of the US population and to correct for nonresponse bias, and all standard errors were calculated using replicate weights. For the analysis of health information brokering, we conducted cross-tabulation and chi-square statistics to evaluate the relationship between brokering and sociodemographic characteristics, and between brokering and online information seeking experiences. Additionally, we performed cross-tabulation and chi-square statistics on the following subgroups: those with income <$20,000 ($n$=744) and those born outside the United States (n=508), again examining the relationship between brokering and sociodemographic characteristics and online information seeking behaviors and experiences. All chi-square statistics had a Rao–Scott correction to account for the complex nature of the survey. A multivariable modified Poisson regression analysis was conducted examining the independent associations of age, sex, race or ethnicity, household income (imputed), and education with health information brokering. Modified Poisson regression was used because in cases of high prevalence, the prevalence ratio (PR) is a better approximation to the relative risk than the odds ratio [24]. Curators of HINTS data use Cox-Iannacchione weighted sequential hot deck imputation to impute values for missing income data.

Results

Our final sample included 3142 respondents. Approximately half (54.8 **%**) of the respondents reported acting as health information brokers.

Sociodemographic Characteristics of Brokers (Brokers vs Nonbrokers)

On bivariate analyses (Table 1), brokers were more frequently female (58.4% of brokers were women vs 43.8% of nonbrokers; P<.001). Brokers reported higher incomes (36.0% earned ≥ $75,000 per year vs 29.0% of nonbrokers; P=.02) and higher educational levels (73.6% had at least some college degree vs 56.8% of nonbrokers; P<.001). Respondents between the ages 35 and 64 acted as brokers most frequently (32.7% of health information brokers were aged 35-49 vs 27.7% of nonbrokers; 26.5% of brokers were aged 50-64 vs 23.5% of nonbrokers, P=.01). Compared to nonbrokers, a higher percentage of brokers were married (61.6% of brokers vs 55.1% of nonbrokers; P=.036).

Table 1. Sociodemographic characteristics of health information brokers[a]

Characteristic		Total	Health information broker	Not health information broker	P value[c]
		N (%)[b]	N (%)[b]	N (%)[b]	
Overall		3142 (100)	1774 (54.8)	1368 (45.2)	
Age					.013
	18-34	422 (27.22)	239 (25.97)	183 (28.74)	
	35-49	701 (30.39)	428 (32.65)	273 (27.65)	
	50-64	1060 (25.13)	628 (26.49)	432 (23.49)	
	65-74	504 (9.36)	275 (8.80)	229 (10.04)	
	75+	355 (7.89)	150 (6.09)	205 (10.08)	
Sex					<.001
	Female	1882 (51.77)	1179 (58.37)	703 (43.78)	
	Male	1179 (48.23)	549 (41.63)	630 (56.22)	
Education					<.001
	Less than/High school	977 (33.96)	471 (26.37)	506 (43.18)	
	Some college	924 (32.79)	531 (36.63)	393 (28.12)	
	College	1155 (33.25)	728 (37.00)	427 (28.70)	
Race/Ethnicity					.990
	White, non-Hispanic	1576 (67.27)	859 (67.37)	717 (67.15)	
	Black, non-Hispanic	416 (10.44)	258 (10.35)	158 (10.55)	
	Hispanic	491 (15.02)	291 (14.81)	200 (15.28)	
	Other, non-Hispanic	208 (7.27)	125 (7.48)	83 (7.02)	
Household income					.019
	Less than $20,000	744 (20.75)	402 (20.34)	342 (21.26)	
	$20,000 to < $35,000	437 (13.99)	226 (11.20)	211 (17.42)	
	$35,000 to < $50,000	430 (14.52)	238 (13.82)	192 (15.37)	
	$50,000 to < $75,000	495 (17.87)	285 (18.64)	210 (16.92)	
	$75,000 or more	880 (32.88)	538 (36.00)	342 (29.03)	
Currently employed					.733
	Employed	1600 (61.61)	936 (62.05)	664 (61.08)	
	Not employed	1468 (38.39)	800 (37.95)	668 (38.92)	
Marital status					.036
	Divorced/Separated	603 (11.02)	340 (11.33)	263 (10.63)	
	Married, living as married	1572 (58.68)	936 (61.61)	636 (55.10)	
	Never married	529 (24.52)	288 (22.40)	241 (27.11)	
	Widowed	343 (5.78)	161 (4.66)	182 (7.16)	
Born in the United States					.624
	Yes	2582 (83.95)	1444 (83.44)	1138 (84.57)	
	No	508 (16.05)	303 (16.56)	205 (15.43)	
Speaks English					.067
	Very well	2578 (87.41)	1486 (89.42)	1092 (84.97)	
	Well	269 (7.65)	129 (6.18)	140 (9.44)	
	Not at all, not well	160 (4.94)	94 (4.40)	66 (5.59)	

[a]Health information brokers: respondents to the Health Information National Trends Survey (HINTS 2013-2014) who answer yes to the question: "Do family members and friends ask you for information or advice on health topics?

[b]Percentages are weighted according to the US population estimates in the American Community Survey to provide representative estimates of the adult US population.

[c]Rao–Scott chi-square test, missing excluded

Brokers' Health Information Seeking Experiences

On bivariate analyses (Table 2), brokers more frequently reported having looked for information on health or medical topics from any source (86.2% of brokers had sought health information vs 67.4% of nonbrokers, *P*<.001). Brokers also expressed confidence in their own ability to find information on health and medical topics as needed (58.8% of brokers vs 51.9% of nonbrokers indicated that they felt completely or very confident; *P*=.02). When participants were asked to rate their trust in information about health or medical topics, brokers and nonbrokers did not significantly differ in trust of information derived from the Internet. Brokers less frequently reported trusting their doctor "a lot" compared to nonbrokers (66.1% of brokers vs 72.0% of nonbrokers; *P*=.01).

Brokers more frequently engaged in health information seeking on behalf of others. Almost half of the brokers (48.1%) reported that their most recent information search was on behalf of someone else as compared to 33.8% of non-brokers (*P*<.001).

A lower percentage of brokers reported that they would first consult a health care provider (46.6% of brokers vs 54.8% of nonbrokers; *P*=.021) while many brokers cited the Internet as their first resource for health information when required (44.5% of brokers would use Internet first vs 36.4% of nonbrokers; *P*=.021). While not all surrogate searches were Internet-based, brokers more frequently reported use of the Internet for surrogate health information seeking in the previous 12 months (79.2% of brokers had used the Internet to seek medical information on behalf of someone else vs 52.8% of nonbrokers; *P*<.001 (see Table 3).

Internet Experience

Health information brokers demonstrated greater use of the Internet (84.4% of brokers had used the Internet vs 71.4% of nonbrokers; *P*<.001). Once online, brokers more frequently pursued a number of health information seeking activities. Compared to nonbrokers, a higher proportion of brokers used the Internet to look for a health care provider, to download

health information, and to track personal health information (Table 3). Health-related YouTube videos had been viewed by 42.3% of brokers as compared to only 25.2% of nonbrokers (*P*<.001). While both brokers and nonbrokers visited social networking sites (77.4% of brokers vs 74.3% of nonbrokers; *P* =NS), brokers shared health information on such sites more frequently (30.0% of brokers vs 14.5% of nonbrokers; *P*<.001). Brokers frequently participated in an online forums or support groups for people with similar health or medical issues (8.6% of brokers vs 4.7% of nonbrokers; *P*=.034) and communicated online with a doctor or someone in the doctor's office (36.3% of brokers vs 20.4% of nonbrokers; *P*<.001).

Low-Income Health Information Brokers

Among respondents earning less than $20,000 annually (*n*=744), 54.1% were health information brokers. Education level was the only demographic variable associated with health information brokering among low-income respondents; compared to those with high school or less education, those with greater than high school education more frequently acted as brokers (53.4% of brokers vs 38.9% of nonbrokers; *P*=.0138). No other demographic variables were significantly associated with brokering among those with annual incomes less than $20,000.

Additionally, among those with low incomes, rates of seeking health information (from any source) were higher for brokers (81.1% vs 56.3% of nonbrokers; *P*<.001) and brokers were more often engaged in online surrogate seeking (72.6% vs 44.1%; *P*=.007). Low-income brokers more frequently reported trust in the Internet (15.8% had 'a lot' of trust in information from the Internet vs. 5.0% of non-brokers; *P*=.02). Brokers in this low-income group (compared to nonbrokers) used the Internet more often (see Figure 1) and were more likely to have participated in a social networking site, downloaded health information and used the Internet to communicate with doctors or personnel in doctors' offices. We also observed a trend toward more frequent online tracking of personal health information among brokers (25.5% of brokers vs 11.6%; *P*=.06).

Table 2. Health information seeking experiences of health information brokers[a]

Experience	Total N (%)[b]	Health information Broker N (%)[b]	Not health information Broker N (%)[b]	P-value[c]
Ever looked for health information (any source)?				<.001
Yes	2482 (77.72)	1528 (86.20)	954 (67.43)	
No	660 (22.28)	246 (13.80)	414 (32.57)	
Most recent searched, sought health information for whom?				<.001
Self	1466 (57.51)	820 (51.92)	646 (66.22)	
Someone else	409 (18.70)	272 (20.28)	137 (16.26)	
Both self and someone else	587 (23.78)	426 (27.80)	161 (17.52)	
Confidence in ability to obtain needed health information				.020
Complete/Very	1779 (55.72)	1059 (58.81)	720 (51.90)	
Somewhat/A little/Not at all	1292 (44.28)	688 (41.19)	604 (48.10)	
Where did you seek health information most recently?				.362
Family/Friends	92 (4.12)	56 (4.18)	36 (4.03)	
Health care provider	400 (14.77)	229 (13.03)	171 (17.55)	
Internet	1328 (69.50)	824 (71.37)	504 (66.53)	
Print material	241 (9.46)	144 (9.18)	97 (9.91)	
Other	62 (2.14)	39 (2.24)	23 (1.97)	
If strongly needed, where would you seek health information?				.021
Family/Friends	136 (4.37)	65 (3.80)	71 (5.07)	
Health care provider	1625 (50.25)	863 (46.55)	762 (54.75)	
Internet	1043 (40.82)	647 (44.46)	396 (36.39)	
Print material	99 (2.69)	51 (3.10)	48 (2.18)	
Other	69 (1.87)	39 (2.09)	30 (1.61)	
Trust information on health or medical topics from doctor				.014
A lot	2135 (68.77)	1188 (66.12)	947 (71.99)	
Some/A little/Not at all	961 (31.23)	564 (33.88)	397 (28.01)	
Trust information on health or medical topics from family or friends				.185
A lot	230 (8.40)	140 (7.20)	90 (9.86)	
Some/A little/Not at all	2773 (91.60)	1570 (92.80)	1203 (90.14)	
Trust information on health or medical topics from news				.204
A lot	688 (24.15)	418 (25.50)	270 (22.51)	
Some/A little/Not at all	2329 (75.85)	1300 (74.50)	1029 (77.49)	
Trust information on health or medical topics from radio				.056
A lot	68 (1.74)	35 (1.11)	33 (2.51)	
Some/A little/Not at all	2883 (98.26)	1642 (98.89)	1241 (97.49)	
Trust information on health or medical topics from the Internet				.056
A lot	417 (12.31)	268 (14.04)	149 (10.18)	
Some/A little/Not at all	2530 (87.69)	1412 (85.96)	1118 (89.82)	
Trust information on health or medical topics from television				.551
A lot	237 (7.88)	148 (7.35)	89 (8.52)	
Some/A little/Not at all	2771 (92.12)	1564 (92.65)	1207 (91.48)	

XSL•FO
RenderX

Experience	Total	Health information Broker	Not health information Broker	P-value[c]
	N (%)[b]	N (%)[b]	N (%)[b]	
Trust information on health or medical topics from government				.424
A lot	765 (26.58)	462 (27.54)	303 (25.41)	
All others	2204 (73.42)	1230 (72.46)	974 (74.59)	
Trust information on health or medical topics from charities or religious organizations				.367
A lot	296 (9.46)	185 (10.14)	111 (8.63)	
All others	2700 (90.54)	1523 (89.86)	1177 (91.37)	

[a]Health information brokers: respondents to the Health Information National Trends Survey (HINTS 2013-2014) who answer yes to the question: "Do family members and friends ask you for information or advice on health topics?

[b]Percentages are weighted according to the US population estimates in the American Community Survey to provide representative estimates of the adult US population.

[c]Rao–Scott chi-square test, missing excluded

Figure 1. Comparison of health information brokers versus nonbrokers within low-income respondents (<$20,000 annually; n=744).
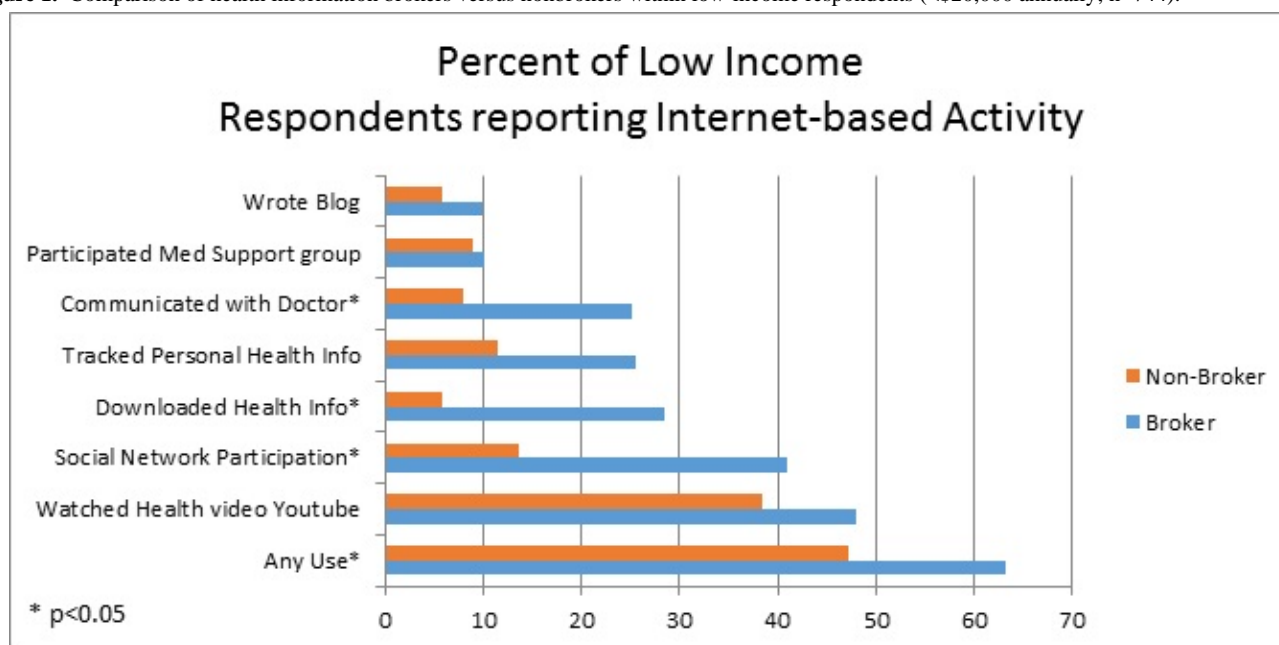
Health Information Brokers Born Outside the United States

Subgroup analyses restricted to respondents born outside of the United States (*n*=508) show that 56.5% of them were brokers. As with low-income brokers, education was the only sociodemographic variable significantly associated with broker status among foreign-born respondents (69.5 % of brokers vs 47.1% of nonbrokers had completed education past high school; *P*=.0012).

Similar to the overall population and to the low-income subgroup, rates of seeking health information (from any source) were higher for foreign-born brokers than nonbrokers (81.9% vs 51.5% respectively; *P*<.001) and foreign-born brokers more often sought health information on the Internet for someone else (81.5% vs 50.2%; *P*<.001). Foreign-born brokers showed no significant difference in trusting the Internet but less often trusted health or medical information from television (10.4% vs 25.8% of foreign-born nonbrokers; *P*=.007) or from charities or religious organizations (10.7% vs 24.1%; *P*=.004); they selected the Internet as a preferred source of information more frequently (45.6% vs 31.1% of nonbrokers; *P*=.001). As shown in Figure 2 , a higher proportion of these brokers reported some form of Internet use. Compared to foreign born nonbrokers, these brokers were more likely to have written in an online diary or blog, participated in an online medical support group, watched a health video on YouTube, downloaded health information, and tracked personal health information online. We found a trend toward increased frequency of Internet-based communication with a personnel in the doctor's office (45.1% of brokers vs 27.1% of nonbrokers; *P*=.09).

Table 3. Internet experiences and online health information exchange preferences expressed by health information brokers[a]

Experience	Total	Health information broker	Not health information broker	P-value[c]
	N (%)[b]	N (%)[b]	N (%)[b]	
Used Internet				<.001
Yes	2266 (78.51)	1393 (84.36)	873 (71.39)	
No	858 (21.49)	373 (15.64)	485 (28.61)	
Visited social networking site such as Facebook or LinkedIn [d]				.334
Yes	1625 (76.14)	1012 (77.41)	613 (74.33)	
No	628 (23.86)	369 (22.59)	259 (25.67)	
Shared health information on social networking site such as Facebook or Twitter [d]				<.001
Yes	446 (23.64)	333 (30.02)	113 (14.53)	
No	1803 (76.36)	1045 (69.98)	758 (85.47)	
Wrote in online diary or blog [d]				.176
Yes	138 (6.35)	91 (7.24)	47 (5.06)	
No	2109 (93.65)	1286 (92.76)	823 (94.94)	
Participated in online forum or support group for people with similar health or medical issues [d]				.034
Yes	150 (7.02)	115 (8.61)	35 (4.74)	
No	2097 (92.98)	1263 (91.39)	834 (95.26)	
Watched health-related video on YouTube [d]				<.001
Yes	702 (35.30)	508 (42.34)	194 (25.23)	
No	1543 (64.70)	867 (57.66)	676 (74.77)	
Used the Internet to				
Look for medical information for yourself [d]				<.001
Yes	1808 (79.66)	1175 (85.41)	633 (71.47)	
No	439 (20.34)	200 (14.59)	239 (28.53)	
Look for medical information for someone else [d]				<.001
Yes	1465 (68.28)	1038 (79.15)	427 (52.75)	
No	775 (31.72)	333 (20.85)	442 (47.25)	
Look for information on quitting smoking				.284
Yes	159 (9.13)	108 (8.27)	51 (10.37)	
No	2079 (90.87)	1262 (91.73)	817 (89.63)	
Buy medicine or vitamins				.007
Yes	456 (20.30)	302 (23.18)	154 (16.18)	
No	1784 (79.70)	1070 (76.82)	714 (83.82)	
Look for a healthcare provider				.004
Yes	826 (38.85)	560 (43.40)	266 (32.36)	
No	1407 (61.15)	803 (56.60)	604 (67.64)	
Download health information to mobile device				<.001
Yes	432 (20.02)	341 (27.43)	91 (9.43)	
No	1807 (79.98)	1031 (72.57)	776 (90.57)	
Track personal health information (eg, care received, test results, medical appointments)				.001
Yes	698 (28.29)	497 (33.64)	201 (20.66)	

XSL•FO
RenderX

Experience	Total	Health information broker	Not health information broker	P -value[c]
	N (%)[b]	N (%)[b]	N (%)[b]	
No	1541 (71.71)	873 (66.36)	668 (79.34)	
Communicate with doctor or personnel in doctor's office				<.001
Yes	666 (29.74)	458 (36.32)	208 (20.35)	
No	1576 (70.26)	914 (63.68)	662 (79.65)	

[a]Health information brokers: respondents to the Health Information National Trends Survey (2013-2014) who answered yes to the question: "Do family members and friends ask you for information or advice on health topics?"

[b]Percentages are weighted according to the US population estimates in the American Community Survey to provide representative estimates of the adult US population.

[c]Rao–Scott chi-square test, missing excluded

[d]In the last 12 months

Figure 2. Comparison of health information brokers versus nonbrokers among respondents born outside the United States (n=508).
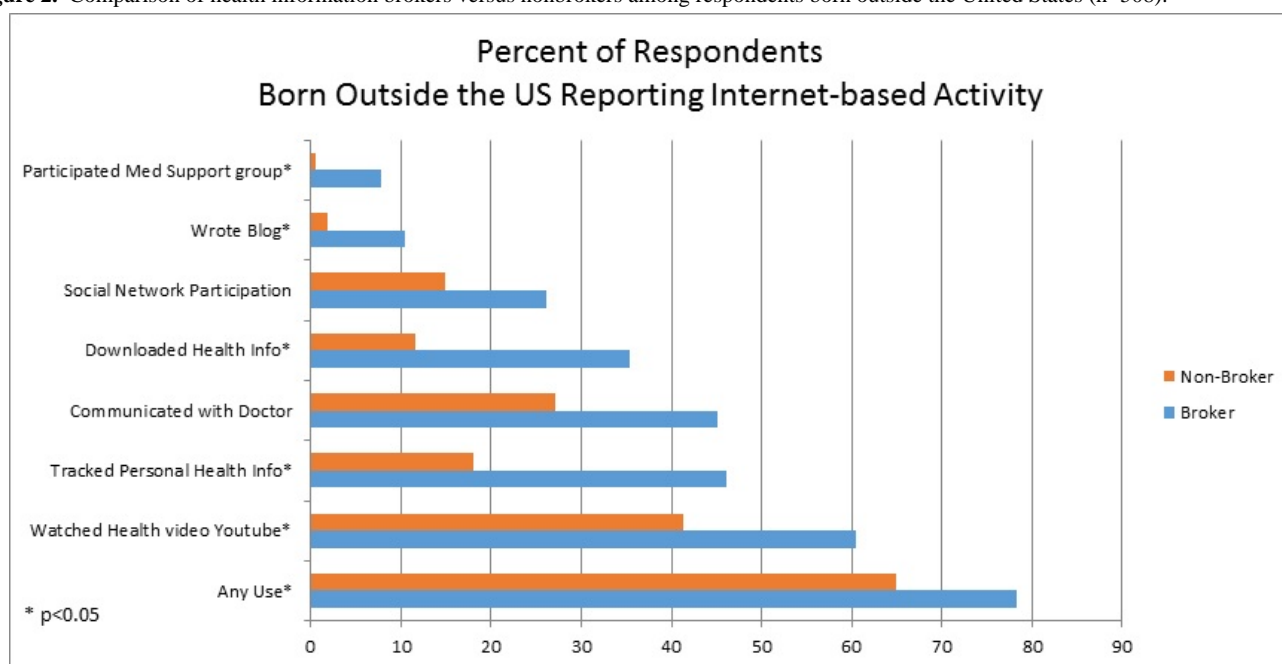
Multivariable Models

Multivariable models based on the entire sample (Table 4) predicting broker status from age, sex, race or ethnicity, income level, and educational level found that those whose education was greater than high school were significantly more likely to have acted as brokers (PR 1.42, 95% CI, 1.22-1.65) compared to those with education of high school or less. Women were significantly more likely to report acting as brokers compared to men (PR 1.34, 95% CI, 1.23-1.47) and elderly respondents aged 75 and above were significantly less likely to report acting as brokers when compared with those aged 35-49 years (PR 0.81, 95% CI, 0.67-0.99).

Table 4. Multivariate analysis.

Characteristics		Predicting information broker status	
		Prevalence ratio (PR)	95% CI
Age (35-49)			
	18-34	0.85	0.72-1.01
	50-64	0.94	0.84-1.07
	65-74	0.90	0.78-1.05
	75+	0.81[a]	0.67-0.99
Female		1.34[b]	1.23-1.47
Race (White, non-Hispanic)			
	Black, non-Hispanic	0.99	0.84-1.17
	Asian, non-Hispanic	0.99	0.75-1.30
	Hispanic	1.13	0.94-1.35
	Other/Unknown	1.05	0.86-1.28
Household income imputed ($75,000 or more)			
	Less than $20,000	0.97	0.80-1.18
	$20,000-< $35,000	0.82	0.65-1.04
	$35,000-< $50,000	0.92	0.79-1.07
	$50,000-< $75,000	1.01	0.87-1.17
Education (≤high school)			
	Above high school	1.42[b]	1.22-1.65
Observations		2870	

[a]$P<.05$

[b]$P<.001$

Discussion

Principal Findings

To our knowledge, this is the first study to examine characteristics and online behaviors of health information brokers within a national population. Our study builds on previous research, highlighting the widespread nature of health information brokering and documenting characteristics of brokers. As a group, health information brokers were more often female and had higher educational levels. Although higher rates of brokering were seen amongst those in the highest income groups, we found this behavior to be present across all income levels. We also found comparable brokering rates among those born in and those born outside the United States. Our findings add to the literature by providing a deeper understanding of health information brokers' online behavior. Brokers used the Internet more often for a variety of tasks associated with health information-seeking and information exchange. They also sought information on behalf of others more frequently and reported greater confidence in their own ability to obtain needed health information. Finally, this study offers insight into the existence of brokering activity among those who do not view themselves as active seekers of information. While the majority of brokers described themselves as health information seekers, more than

1 out of 10 brokers did not. This subset of brokers merits further investigation to understand whether they view themselves as sources of health information or advice based on their own experiences as opposed to their skill in information seeking.

We found over half of the respondents had engaged in health information brokering activity; this documentation of the widespread use of interpersonal sources of health information is consistent with previous studies [5,25]. According to a 2014 Pew study, when faced with a serious health issue, the majority of adults surveyed (60%) turned to friends and family for health information or support; one-quarter of those surveyed (24%) sought information or support from patients with the same health condition and 13% of adults conducted online searches to find others with health concerns similar to theirs, a practice more common among those with chronic and rare medical conditions [5]. In a survey assessing cancer screening decision-making among a randomly selected national sample of adults 50 years and above [2], authors found that family and friends were frequently ranked as an extremely important source of information for prostate, colorectal, and breast cancer screening, second only to health care providers. Studies in the US estimate that 56-66% of those responding to national surveys have sought health information on behalf of others [13,14]. In contrast, a 2011 online survey administered to online health information

seekers in the United Kingdom who were accessing the National Health Service direct website showed that only 30% sought information for others [26].

Individuals possessing knowledge on a health topic (whether derived from searches or personal experience) and who consider it important to share this information with others have been described previously as "health information mavens." Kontos et al [18] found that being female was associated with scoring higher on a 5-item scale measuring "health information mavenism"; similarly, we found that health information brokers were more likely to be female. Our study adds to the current understanding of health information brokering activity, documenting an association between higher education and brokering. The relation between education and scores was not directly examined by Kontos et al, but participants had both a lower frequency of education past the high school level (7% compared with 76% in our study) and an overall lower percentage of respondents self-identifying as mavens (44%). Women and those with higher education have been identified previously as more likely to engage in seeking health information on behalf of others [13,14].

While we did not find a significant relationship between brokering and either place of birth or language in our national sample, Kontos et al found that those living in the United States for less than 10 years had slightly higher scores than others. In addition, their study found that those who did not speak English as a primary language but read and spoke English occasionally at home were more likely to have higher scores than those who had English as their primary language. For more recent immigrants and for those with developing English skills, Kontos et al postulated that being part of an interdependent community might have a protective effect, adding that the ability to "effectively communicate in both English and their native language" likely made these individuals "conduits of information, including health information, for their community" [18].

Our understanding of the online behaviors of health information brokers indicates that they are more active than the nonbrokers in online communities. Our study shows that these individuals are more likely to create and exchange online content, whether by participating in social networks and medical support groups or by communicating with personnel in doctors' offices, and they are more likely to seek, download, and track health information. Similar online behaviors have been identified previously among those who seek health information on behalf of others [13,14].

While the "digital divide" has narrowed in recent years, there remain disparities in Internet use and access. These disparities disproportionately affect the elderly and those who are less affluent, have less education or who live in rural parts of the country [27,28]. Brokers may help decrease the impact of this divide by accessing and passing along health information; however, more research is needed to understand the quality of the information disseminated by brokers. Commonly used search engines such as Google do not retrieve health information based on the quality (or appropriateness for a given health literacy level) and information shared through online communities and support groups is not always accurate. This places the onus on the broker to judge the quality and relevance of the information before sharing it with others. Additional insights into the attitudes attitudes and behaviors of brokers may be useful for improving Internet-based dissemination of high-quality health information.

Based on our findings, further exploration of ways in which brokers support patient engagement and disease management is encouraged. Our findings indicate that brokers are more likely to have communicated online with doctors or personnel in the doctor's office than nonbrokers, but it is not clear whether brokers are doing this on behalf of themselves or for someone else (for instance, through secure messaging via proxy portal access within an electronic health record, EHR). Future studies might examine the current use and potential benefits incurred by granting brokers (with permission from patients) proxy access to patient EHRs. Such access may help brokers who are already partnering with patients in acute illness and in chronic disease self-management efforts. Family or friends who accompany patients during healthcare visits could also be directed to approved online resources such as the Centers for Disease Control and Prevention (CDC) caregiving resources (http://www.cdc.gov/aging/caregiving/resources.htm). This can be accomplished via hyperlinks in the patient portal and could facilitate access to quality information on behalf of patients in the ambulatory setting, and (where available) on behalf of hospitalized patients (via applications such as Epic MyChart Bedside). Additional functions on technology-based interventions could be developed to support proxy access and communication of information to others. Family or friends could also be provided information on recommended websites via printed care summaries, which are routinely provided to patients at the end of ambulatory visits.

Studies examining the link between social network characteristics and broker behavior would help us understand whether health communication campaigns would benefit from targeting health information brokers. Such studies could also examine whether being a health information broker is a static characteristic or a behavior which changes over time.

Finally, our study highlights the need for further investigation into brokering among those who do not view themselves as active seekers of information; in particular, brokers who dispense information and advice based on personal experience with the medical system.

Limitations

Some limitations of this research are worth noting. HINTS data are cross-sectional; therefore causality cannot be inferred. Response rates for HINTS, while consistent with other national surveys, were low [29]. National surveys, such as HINTS, are very often constrained by survey length and respondent burden. Therefore, the number of items used to measure a multifaceted behavior, such as information seeking on behalf of others, may not fully capture the constructs of interest. Use of self-report data introduces the possibility of recall bias. Finally, it is important to recognize that lay sources of information may not always transmit medically accurate or guideline-concordant information. This study did not assess health literacy level and

did not explore the content or quality of health information or advice transmitted.

Conclusion

In a national sample, a high proportion of respondents, including those in the lowest income levels and those born outside the US, report acting as brokers of health information. Brokers more frequently engaged in a variety of online behaviors including health information seeking, creation of online content and downloading of health information onto a mobile device. Members of traditionally disadvantaged groups who report acting as health information brokers display these same behaviors and future studies should examine whether and how these brokers are narrowing the existing digital divide. Directing

brokers to high-quality Internet-based resources in familiar online venues or to resources designed for downloading may be an effective way to support dissemination of health information.

Practice Implications

Health information brokers have important lessons to teach healthcare professionals about the role they play in disseminating health information and advice to their friends and families. These brokers may be effective targets for initiatives aimed at supporting patient engagement and disease management. In addition, self-identification as a person who engages in brokering behavior may be a marker for those well-positioned to assist in health communication campaigns.

Conflicts of Interest

None declared.

References

1. Fox S. Pew Research Center; updated January 15. 2014. The social life of health information URL: http://www.pewresearch.org/fact-tank/2014/01/15/the-social-life-of-health-information/ [accessed 2016-05-01] [WebCite Cache ID 6hC8r8b8f]
2. Hoffman RM, Lewis CL, Pignone MP, Couper MP, Barry MJ, Elmore JG, et al. Decision-making processes for breast, colorectal, and prostate cancer screening: the DECISIONS survey. Med Decis Making 2010;30(5 Suppl):53S-64S [FREE Full text] [doi: 10.1177/0272989X10378701] [Medline: 20881154]
3. Kjos A, Worley M, Schommer J. University of Minnesota Digital Conservancy.: University of Minnesota, College of Pharmacy; 2011. Medication information seeking behavior in a social context: the role of lay and professional social network contacts URL: http://conservancy.umn.edu/handle/11299/120061 [accessed 2016-05-01] [WebCite Cache ID 6hC9Prtya]
4. Morey OT. Health information ties: preliminary findings on the health information seeking behaviour of an African-American community. Information Research 2006;12(2):9.
5. Pew Research Center. 2014 Jun 19. Health Fact Sheet URL: http://www.pewinternet.org/fact-sheets/health-fact-sheet/ [accessed 2016-05-01] [WebCite Cache ID 6hC9lWZ8F]
6. Gross M. The imposed query. RQ. 1995. (2) p. 236-243 URL: http://www.jstor.org/stable/20862879 [WebCite Cache ID 6hYXb1PNl]
7. Gross M. Imposed information seeking in public libraries and school library media centres: a common behaviour. Information research 2001;6(2):6-2 [FREE Full text]
8. Gross M, Saxton ML. Who Wants To Know? Imposed Queries in the Public Library. Public Libraries. 2001. (3) p. 170-176
9. Metoyer-Duran C. Information-seeking behavior of gatekeepers in ethnolinguistic communities: Overview of a taxonomy. Library & Information Science Research 1991;13(4):319-346.
10. Rioux KS. Information acquiring-and-sharing in Internet-based environments: An exploratory study of individual user behaviors. 2004. URL: https://repositories.lib.utexas.edu/bitstream/handle/2152/2173/riouxks042.pdf [accessed 2016-05-02] [WebCite Cache ID 6hCnCNGBp]
11. Abrahamson JA, Fisher KE. 'What's past is prologue': towards a general model of lay information mediary behaviour. Information Research 2007;12:1-21.
12. Cutrona SL, Roblin DW, Wagner JL, Gaglio B, Williams AE, Torres SR, et al. Adult Willingness to Use Email and Social Media for Peer-to-Peer Cancer Screening Communication: Quantitative Interview Study. JMIR Res Protoc 2013;2(2):e52 [FREE Full text] [doi: 10.2196/resprot.2886] [Medline: 24287495]
13. Sadasivam RS, Kinney RL, Lemon SC, Shimada SL, Allison JJ, Houston TK. Internet health information seeking is a team sport: analysis of the Pew Internet Survey. Int J Med Inform 2013 Mar;82(3):193-200. [doi: 10.1016/j.ijmedinf.2012.09.008] [Medline: 23149121]

XSL·FO

RenderX

14.   Cutrona SL, Mazor KM, Vieux SN, Luger TM, Volkman JE, Finney Rutten Lila J. Health information-seeking on behalf of others: characteristics of "surrogate seekers". J Cancer Educ 2015 Mar;30(1):12-19 [FREE Full text] [doi: 10.1007/s13187-014-0701-3] [Medline: 24989816]

15.   Abrahamson JA, Fisher KE, Turner AG, Durrance JC, Turner TC. Lay information mediary behavior uncovered: exploring how nonprofessionals seek health information for themselves and others online. J Med Libr Assoc 2008 Oct;96(4):310-323 [FREE Full text] [doi: 10.3163/1536-5050.96.4.006] [Medline: 18974809]

16.   Kontos E, Blake KD, Chou WS, Prestin A. Predictors of eHealth usage: insights on the digital divide from the Health Information National Trends Survey 2012. J Med Internet Res 2014;16(7):e172 [FREE Full text] [doi: 10.2196/jmir.3117] [Medline: 25048379]

17.   Cialdini RB, Brown SL, Lewis BP, Luce C, Neuberg SL. Reinterpreting the empathy-altruism relationship: when one into one equals oneness. J Pers Soc Psychol 1997 Sep;73(3):481-494. [Medline: 9294898]

18.   Kontos EZ, Emmons KM, Puleo E, Viswanath K. Determinants and beliefs of health information mavens among a lower-socioeconomic position and minority population. Soc Sci Med 2011 Jul;73(1):22-32 [FREE Full text] [doi: 10.1016/j.socscimed.2011.04.024] [Medline: 21683493]

19.   Nelson DE, Kreps GL, Hesse BW, Croyle RT, Willis G, Arora NK, et al. The Health Information National Trends Survey (HINTS): development, design, and dissemination. J Health Commun 2004;9(5):443-60; discussion 81. [doi: 10.1080/10810730490504233] [Medline: 15513791]

20.   National CI. National Cancer Institute. Final Report: Health Information National Trends Survey Instruments URL: http:/ /hints.cancer.gov/instrument.aspx [accessed 2016-05-02] [WebCite Cache ID 6hCnYg0gl]

21.   Westat. National Cancer Institute. 2014. Health Information National Trends Survey 4 (HINTS 4) Cycle 3 Methodology Report URL: http://hints.cancer.gov/docs/Instruments/HINTS_4_Cycle_3_Methods_Report_FINAL_508c_03_21_2014. pdf [accessed 2016-05-02] [WebCite Cache ID 6hCnb3dCr]

22.   Lumley T. Analysis of complex survey samples. Journal of Statistical Software 2004;9(1):1-19.

23.   SAS Institute, Inc. 2011. SAS/STAT Software, Version 9 URL: http://www.sas.com/en_us/home.html [accessed 2016-05-02] [WebCite Cache ID 6hCnsXJvl]

24.   Zou G. A modified poisson regression approach to prospective studies with binary data. Am J Epidemiol 2004 Apr 1;159(7):702-706 [FREE Full text] [Medline: 15033648]

25.   Larkey LK, Hecht ML, Miller K, Alatorre C. Hispanic cultural norms for health-seeking behaviors in the face of symptoms. Health Educ Behav 2001 Feb;28(1):65-80. [Medline: 11213143]

26.   Powell J, Inglis N, Ronnie J, Large S. The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study. J Med Internet Res 2011;13(1):e20 [FREE Full text] [doi: 10.2196/jmir.1600] [Medline: 21345783]

27.   Council of Economic Advisers. Council of Economic Advisers Reports and Issue Briefs. 2015 Jul. Council of Economic Advisers Issue Brief URL: https://www.whitehouse.gov/sites/default/files/wh_digital_divide_issue_brief.pdf [accessed 2016-05-02] [WebCite Cache ID 6hCoSvjpy]

28.   Perrin AM. Pew Research Center. 2015. Americans' Internet Access: 2000-2015 URL: http://www.pewinternet.org/2015/ 06/26/americans-internet-access-2000-2015/ [accessed 2016-05-02] [WebCite Cache ID 6hCoWtBFW]

29.   Fahimi M, Link M, Mokdad A, Schwartz DA, Levy P. Tracking chronic disease and risk behavior prevalence as survey participation declines: statistics from the behavioral risk factor surveillance system and other national surveys. Prev Chronic Dis 2008 Jul;5(3):A80 [FREE Full text] [Medline: 18558030]

Abbreviations

EHR: Electronic health record
HINTS: Health Information National Trends Survey
PR: Prevalence ratio

XSL•FO
RenderX

Original Paper

Benefits of Diabetes Self-Management for Health Plan Members: A 6-Month Translation Study

Kate Lorig[1], Dr PH; Philip L Ritter[1], PhD; Ralph M Turner[2], PhD; Kathleen English[3], MBA; Diana D Laurent[1], MPH; Jay Greenberg[4], ScD

[1]Stanford School of Medicine, Stanford University, Palo Alto, CA, United States
[2]HealthCore, Wilminton, DE, United States
[3]Anthem, Inc, Indianapolis, IN, United States
[4]NCOA Services, LLC, National Council on Aging, Arlington, VA, United States

Corresponding Author:
Philip L Ritter, PhD
Stanford School of Medicine
Stanford University
1000 Welch Rd., Suite 204
Palo Alto, CA, 94304
United States
Phone: 1 650 725 2873
Fax: 1 650 725 9422
Email: philr@stanford.edu

Abstract

Background: Diabetes self-management education has been shown to be effective in controlled trials. However, few programs that meet American Association of Diabetes Educators standards have been translated into widespread practice.

Objective: This study examined the translation of the evidence-based Better Choices, Better Health-Diabetes program in both Internet and face-to-face versions.

Methods: We administered the Internet program nationally in the United States (n=1010). We conducted face-to-face workshops in Atlanta, Georgia; Indianapolis, Indiana; and St. Louis, Missouri (n=232). Self-report questionnaires collected health indicator, health behavior, and health care utilization measures. Questionnaires were administered on the Web or by mail. We determined hemoglobin A_{1c} (HbA_{1c}) from blood samples collected via mailed kits. Paired t tests determined whether changes between baseline and 6 months differed significantly from no change. Subgroup analyses determined whether participants with specific conditions benefited (high HbA_{1c}, depression, hypoglycemia, nonadherence to medication taking, and no aerobic exercise). We calculated the percentage of participants with improvements of at least 0.4 effect size in at least one of the 5 above measures.

Results: Of the 1242 participants, 884 provided 6-month follow-up questionnaires. There were statistically significant improvements in 6 of 7 health indicators (including HbA_{1c}) and in 7 of 7 behaviors. For each of the 5 conditions, there were significant improvements among those with the condition (effect sizes 0.59–1.1). A total of 662 (75.0%) of study participants improved at least 0.4 effect size in at least one criterion, and 327 (37.1%) improved in 2 or more.

Conclusions: The Diabetes Self-Management Program, offered in two modes, was successfully disseminated to a heterogeneous national population of members of either insured or administered health plans. Participants had small but significant benefits in multiple measures. The program appears effective in improving diabetes management.

XSL•FO
RenderX

Introduction

Background

Type 2 diabetes is a growing problem. Not only is the prevalence increasing, but also there are increased costs and decreased quality of life. Patients, health care providers, families, employers, and insurers all share this burden. The US government, in Healthy People 2020, set national objectives for improvements in diabetes care and outcomes [1].

Diabetes self-management education has been shown to be effective in improving health behaviors and in some cases lowering hemoglobin A_{1c} (HbA$_{1c}$). However, <7% of privately insured persons receive formal education in the year after diagnosis [2] and <60% report having attended a diabetes class [1]. The American Association of Diabetes Educators has set standards for diabetes self-management education [3,4]. Few programs, however, that meet these standards have been translated into widespread practice and shown to be effective. One reason for this is that most programs are not easily translatable: they do not have (1) standardized curricula that are manualized in such a way that minute-to-minute details are given of exact content to be presented, as well as the processes by which this content is offered, (2) standardized training programs, and (3) manuals for fidelity standards. All 3 components are necessary for effective replication.

The Administration for Community Living (US Department of Health and Human Services) and the Evidence-Based Leadership Council have varied in their definitions of evidence-based programs [5,6]. However, key to both definitions are that the program must have been shown to be effective and that it have one or more peer reviewed publications. In addition, there must be infrastructure for translation, and the program must have been replicated in sites other than where it originated. The National Diabetes Education Program does not define evidence-based diabetes education but does suggest using programs that meet the National Standards for Diabetes Self-Management Education [3]. None of these definitions include the notion that the intervention must be shown to be effective outside of its home site.

A National Translation Study

In this paper, we present the national translation of an evidence-based diabetes education program, The Stanford Diabetes Self-Management Program, also known as Better Choices, Better Health-Diabetes (BCBH-D) [7-9], given via the Internet or in small face-to-face community groups.

Recent systematic reviews have found that both community and Internet interventions have been effective in reducing HbA$_{1c}$ and improving quality of life for people with type 2 diabetes [10-14]. These programs have nearly always been offered in the context of a controlled study. In that setting there are often extensive inclusion and exclusion criteria, which may bias the results when applied to the general diabetes population. This is especially true for HbA$_{1c}$, where the inclusion criterion is seldom <7 (considered to be controlled diabetes) and often as high as 9. For example, in Zhang et al's meta-analysis of 20 studies

using peer support for diabetes, 12 studies had a mean baseline HbA$_{1c}$ of ≥8 [14]. There is limited information about the effectiveness of specific programs when translated to widespread use and offered to large populations by non-health care providers [15,16].

For this translation study, the US National Council on Aging led a collaboration of 5 organizations. It also provided the platform for offering the Internet workshops. Anthem, Inc (Indianapolis, IN, USA) -affiliated health plans recruited participants from their members. The Young Men's Christian Association of the United States of America in Atlanta, Georgia, and OASIS Institute in St. Louis, Missouri, and Indianapolis, Indiana, offered community programs. The Stanford Patient Education Research Center (Stanford University School of Medicine, Stanford, CA, USA) collected data and provided data analysis. HealthCore (Wilmington, DE, USA), a subsidiary of Anthem, Inc, contributed to the design and data analysis. The team members met for months prior to the beginning of the study to decide on study design, outcome variables, and logistics. They continued to meet at least monthly for the duration of the intervention. The study was approved by the Stanford University Institutional Review Board and New England Independent Review Board.

We hypothesized that over 6 months, people with diabetes participating in the program would demonstrate (1) a reduction in HbA$_{1c}$, (2) a reduction in symptoms (hypoglycemic symptoms and depression), (3) increases in healthful behaviors (exercise, communication with physicians, and medication adherence), (4) increases in receiving recommended tests (eye, foot, cholesterol, and kidney examinations). We also hypothesized (5) that program effectiveness would be independent of mode of delivery, (6) that participants with baseline HbA$_{1c}$ ≥9.0, 8-item Patient Health Questionnaire (PHQ-8) depression ≥10.0, having 2 or more hypoglycemia symptoms, medication nonadherence, or no aerobic exercise would have clinically significant improvements in the variables of interest, and (7) that most of the participants would have a moderate effect size (0.4) improvement in either reducing HbA$_{1c}$, reducing depression, increasing medication adherence, reducing hypoglycemic symptoms, or increasing exercise.

In addition, we wanted to explore the potential of BCBH-D to meet some of the Healthy People 2020 objectives: decrease the proportion of people with diabetes with HbA$_{1c}$ >9, and increase the proportion of people with diabetes who (1) have an annual foot examination, (2) have an annual urinary microalbumin measurement, and (3) receive formal diabetes education.

Methods

Intervention

We chose BCBH-D because it was developed for people with type 2 diabetes in real-world settings. We built both programs (face-to-face and Internet) on extensive patient input, as well as assistance from certified diabetes educators. Both programs have been shown to be effective in previous randomized trials [7,9], have the same content, and are taught in an interactive manner designed to enhance self-efficacy [17]. Both have a

duration of 6 weeks, 2 peer facilitators, and standardized facilitator training. Participants receive the same book, *Living a Healthy Life with Chronic Conditions* [18], which contains program content and chapters on other chronic conditions. The face-to-face program has detailed facilitator, administrative, and fidelity manuals [19], while the Internet program has an administrative manual for facilitators.

The BCBH-D Internet program is a password-protected, interactive, Web-based program. The user interface consists of 3 major sections. (1) The Learning Center offers 20–30 pages of didactic and interactive content each week. In addition to reading content within the Learning Center, participants make weekly action plans, give feedback on the plan from the previous week, and answer a question such as "What problems do you have with your diabetes?" The action plan, feedback, and questions populate the Discussion Center. (2) The Discussion Center contains 4 interactive bulletin boards: problem solving, action planning, difficult emotions, and celebrations. Participants can post to any of these boards and respond to posts at any time. (3) My Tools is a series of tools, such as a medication log, food diary, exercise diary, and links to other websites that can be used as participants wish.

The Internet facilitators are trained over the Web by first participating in a workshop, followed by Web-based training and then cofacilitating the workshop with an experienced facilitator. Certified diabetes educators are available to facilitators to answer questions as necessary. All interactions between facilitators and participants take place on the Web. Facilitators assist participants with the program. They model action planning and problem solving, and offer encouragement by posting to the discussion boards. Facilitators monitor the daily posts of all participants and report inappropriate posts to program administrators. They also have access to a certified diabetes educator. Unlike in the community program, in the Internet program facilitators do not deliver content, as this is scripted on the Web.

For the Internet program, 25–30 participants log on at least 3 times a week and participate in weekly activities, including reading content, posting an action plan, and interacting on the discussion boards. Any problem a participant wishes to discuss can be posted in the Discussion Center and responded to by other participants. The Internet program mirrors the original community program, except that it does not require real-time attendance. Participants can log in any time, have no face-to-face interaction, and may return to past weeks' material [7,9].

The community workshops meet for 2.5 hours a week; 10–15 people attend and may bring a friend or family member. This is longer than the 10 hours usually covered by insurance participant's health plan benefits. The length was determined by the amount of material to be covered, as well as past work that indicated that shorter programs were less effective [20]. In a recent review, Pillay and colleagues also noted that programs with ≤10 hours of contact had limited benefit in improving glycemic control [21].

Both the community and Internet BCBH-D content meets the diabetes self-management education recommendations for diabetes self-management education and support [3]. More than

20 organizations have received American Association of Diabetes Educators certification for the face-to-face program. Although the content and processes are very similar, there is no means for certifying or recognizing Web-based programs. Both interventions have been widely translated into practice. More than 50,000 people in 39 states in the United States have taken the community program and more than 2000 have participated via the Internet. The community program is also used in 14 countries outside the United States. This broad translation is possible because of program and training standardization, because of the original evidence base of effectiveness, and because program delivery meets the needs of community and health care organizations.

Recruitment

Because we wanted to replicate real-world settings, we stipulated few inclusion (have type 2 diabetes and be covered by an Anthem plan) or exclusion criteria (currently pregnant or had chemotherapy or radiation treatment for cancer in the past year). There were no inclusion criteria based on level of symptom severity, nor was having had cancer by itself an exclusion criterion.

We recruited Internet participants by email or announcements from their employers or emails directly from an Anthem plan. A small percentage of participants were referred by family, friends, or physicians. Participants in both commercial and Medicare Advantage health insurance programs were eligible. Potential study participants went to the recruitment website, completed a screening form, and, if they met study requirements, completed an informed consent and baseline questionnaire.

We recruited face-to-face participants through mailings, flyers in workplaces, case managers, and automated telephone calls. They were asked to call a local number for details about the workshop, screened for study eligibility, and registered. Programs were available in Atlanta, Indianapolis, and St. Louis. A small percentage of the community participants were not covered by an Anthem plan. All other screening criteria were the same for Internet and community participants.

Participants enrolled between October 2013 and October 2014.

Data Collection

We collected data at baseline, 6 months, and 12 months. This paper reports only on the 6-month results. Internet participants completed consent forms and all questionnaires on the Web. Face-to-face participants completed informed consent forms and baseline questionnaires within a week of their first session. Follow-up questionnaires were sent by mail.

We asked potential participants to supply a capillary sample of blood for HbA_{1c} testing. If willing, they were sent a CoreMedica home test kit (CoreMedica Laboratories, Lees Summit, MO, USA). These were returned to the investigators, bar coded to avoid disclosing private information, and then sent to CoreMedica, a Clinical Laboratory Improvement Amendments-certified laboratory [22]. Participants and their physicians were sent the results. Participants were not required to consent to blood testing, nor were they disqualified if they did not return their tests. Because CoreMedica recalibrated its

measurements in June 2014, increasing all values by roughly 0.4, all measures prior to June were adjusted upward by 0.4.

Measures

We chose measures to be of interest to patients, providers, and the health care system. We gathered all data except HbA_{1c} from validated self-report questionnaires. Demographic variables included age, sex, race, ethnicity (non-Hispanic white), education, and marital status. In addition, we recorded other diseases (asthma, chronic obstructive pulmonary disease, other lung disease, hypertension, heart disease, renal disease, arthritis, cancer, depression or other mental condition). Outcome measures, described below and in full detail elsewhere, fell into two broad categories: health indicators and health behaviors.

All health indicators have been validated in previous studies. A higher score indicates greater symptoms or worse health. Self-rated health consists of a single item from the National Health and Nutrition Examination Survey (range 1–5) [23]. The PHQ-8 depression scale is calculated as the sum of 8 items, with a range of 0–24 [24]. The illness intrusiveness scale consists of 13 items that measure how much a participant's illness interferes with different aspects of life [25]. Each item ranges from 1 (not very much) to 7 (very much interference). The hypoglycemic symptoms scale was developed by Piette [26]. It is the; mean of 7 yes–no questions regarding the presence of different symptoms, with a range of 0–7 symptoms. Fatigue and sleep are each single-item visual numeric scales ranging from 0 (no fatigue or no sleep problems) to 10 (severe fatigue or very big problem sleeping) [27].

The health behaviors measures have also been validated elsewhere. A 3-item scale measured communication with physicians. It used a 6-point scale (from never to always) to measure how often the participant prepared a list of questions for the physician related to the illness [28]. We assessed minutes of aerobic exercise per week by asking about 5 types of aerobic exercise [28]. The Morisky Medication Adherence Scale consists of 4 yes–no items asking about taking medication [29]. The scale has a range of 0–4, with a higher value indicating less adherence. We also asked participants if they had eye, foot, cholesterol, and kidney examinations in the past 6 months and past 12 months.

Data Analysis

Translation studies, because of study population heterogeneity, present unique methodological challenges. In many diabetes studies, participants are chosen based on having difficulty with the outcome of interest. Thus, participants enter into studies because of high HbA_{1c}, depression, hypoglycemia, or being nonadherent. In this study, no such screening occurred. This results in greater heterogeneity, as not all participants have the same problems and some may have none of the problems. Consequently, we conducted two types of analyses. The first, or classic, analyses determined the changes, significance, and effect sizes for the population as a whole. The second, or subset, analyses examined only that portion of the study population

that demonstrated problems in the variable of interest, that is, high HbA_{1c} or low adherence to taking medications. A third analysis sought to reconcile these two by examining the percentage of the total population who achieved a moderate benefit (0.4 effect size) in at least one of the variables of interest.

Univariate statistics described demographic characteristics. Independent sample *t* tests compared demographic and baseline outcome variables between those who did not complete 6-month follow-up questionnaires and those who completed them. Paired *t* tests examined changes between baseline and 6 months and whether these differed significantly from a null hypothesis of zero change (hypotheses 1–4).

For participants who had had no examinations (eye, foot, cholesterol, or kidney) in the year prior to entry, we calculated the percentage who had examinations in the 6 months following baseline (hypothesis 4).

To determine effectiveness by mode of delivery (hypothesis 5), we used independent sample *t* tests to compare change scores between Internet and face-to-face participants. We also used analyses of variance to compare differences among the 3 community locations.

We conducted subgroup analyses for participants with specific indicators (hypothesis 6). These were HbA_{1c} >9; clinical depression (PHQ-8 of ≥10 [24]); at least two symptoms of hypoglycemia; low medication adherence; and no exercise at baseline. For each measure, we report the mean change of the group and the percentage who no longer had the negative indication.

To determine whether a large portion of the participants benefited on at least one important outcome, we calculated the percentage who had an improvement of at least 0.4 effect size (change score divided by baseline standard deviation) in at least one of the 5 criterial measures (hypothesis 7). We also calculated the mean number of improvements (out of 5) and examined how this varied by the number of the 5 specific indicators each participant had.

Results

Participants and Baseline Demographics

There were 4639 potential participants who indicated interest in the study. Of these, 1242 participants met the study criteria, consented, completed baseline questionnaires, and attended at least one session of the workshop, resulting in 1010 Internet participants (Figure 1) and 232 face-to-face participants (Figure 2).

Table 1 gives the demographic characteristics. Participants in the Internet group were statistically more likely to be male (*P*=.02, from *t* test), more likely to be married (*P*<.001), less likely to be a racial/ethnic minority (*P*=.001), to be more educated (*P*=.01), and to be younger (*P*<.001). Overall, the participants tended to be well-educated women.

Table 1. Demographic characteristics of the study sample.

Variable	Community (n=232)	Internet (n=1010)	Entire sample (n=1242)
Male, n (%)	61 (26.3%)	346 (34.3%)	407 (32.8%)
Education in years, mean (SD, range)	15.0 (2.86, 8–23)	15.5 (2.75, 9–23)	15.4 (2.78, 8–23)
Married, n (%)	117 (50.4%)	762 (75.5%)	879 (70.8%)
Ethnicity/race, n (%)			
Non-Hispanic white	148 (63.8%)	756 (75.0%)	904 (72.9%)
Black	70 (30.2%)	98 (9.7%)	168 (13.5%)
Hispanic	5 (2.2%)	90 (8.9%)	95 (7.7%)
Age in years, mean (SD, range)	65.6 (10.1, 28–95)	55.0 (8.68, 25–91)	57.0 (9.86, 25–95)
Number of other chronic conditions, mean (SD, range)	1.65 (1.37, 0–8)	1.40 (1.15, 0–6)	1.45 (1.19, 0–8)
Health care coverage, n (%)			
Medicare	133 (57.3%)	62 (6.1%)	195 (15.7%)
Private insurance	166 (71.6%)	1010 (100%)	1176 (94.7%)

Figure 1. Internet group flowchart. HbA_{1c}: hemoglobin A_{ic}.
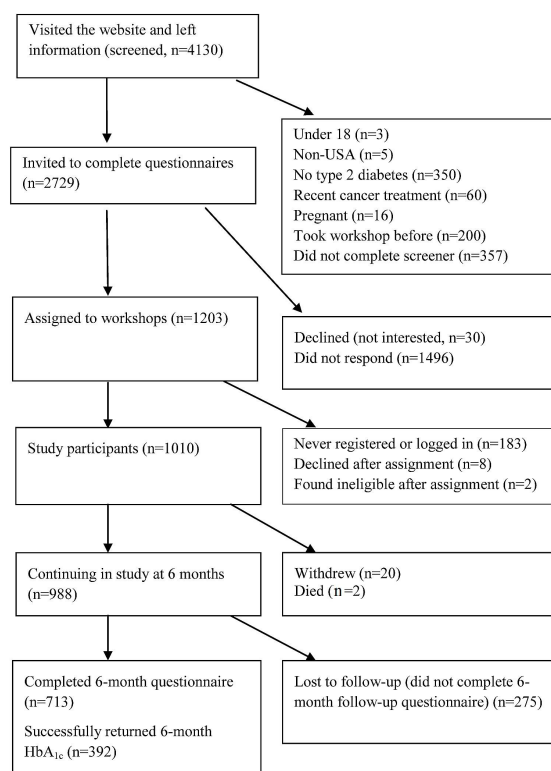
Figure 2. Community workshop group flowchart. HbA_{1c}: hemoglobin A_{1c}.
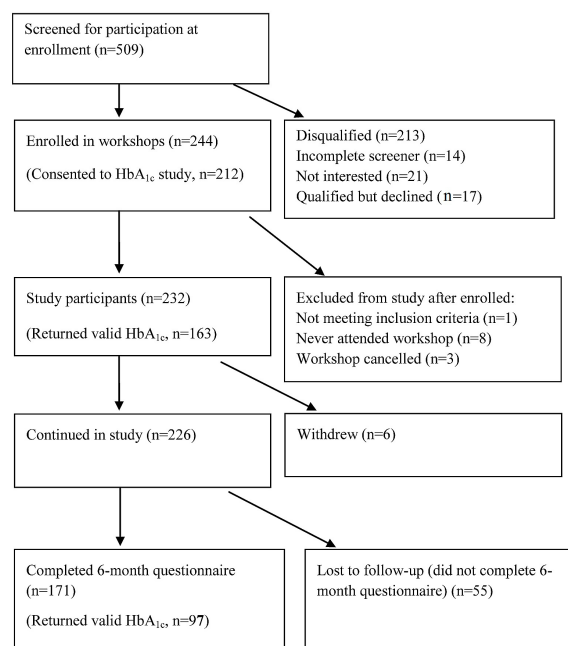
Noncompleters of 6-Month Questionnaires

Outside of HbA_{1c}, which was optional, no variable had more than 1.7% missing data among those who completed questionnaires. We did not attempt to impute missing data.

Of the 1242 participants, 358 (28.82%) did not complete the 6-month questionnaire, including 28 who withdrew from the study or died before 6 months (Figure 1, Figure 2). The study retention rate was similar and nonsignificant by type of program: 713/1010 (70.59%) of the Internet and 171/232 (73.7%) of the community participants.

Comparing 6-month noncompleters and completers on demographic variables, those who did not complete the 6-month follow-up tended to be younger (mean 55.4 vs 57.6 years old, $P<.001$). The groups did not differ by sex, marital status, education, or ethnicity/race. Among 17 baseline outcome variables, 1 had statistically significance differences. Noncompleters reported less medication adherence (1.22 vs 1.06 on a 5-point scale, $P=.02$).

6-Month Changes (Hypotheses 1–4)

Table 2 shows baseline and 6-month change scores for all study participants. It should be noted that at baseline the mean HbA_{1c} was just under 8 with a standard deviation of 1.4. This suggests that, as predicted, the group was very heterogeneous and that many participants had a low HbA_{1c} at baseline.

Table 2. Baseline and 6-month scores for participants with 6-month data in the Internet and workshop groups combined.

Variables	Possible range of values	Desirable score	No.	Baseline		6-Month change		Effect size of change	$P > \| t \|$ for change
				Mean	SD	Mean	95% CI		
Adjusted HbA_{1c}[a,b]		Lower	489	7.95	1.38	–0.151	–2.47 to 10.06	0.109	.002
PHQ-8[c] depression	0–24	Lower	877	5.97	5.02	–1.02	–1.30 to –0.74	0.203	<.001
General health	0–5	Lower	884	2.90	0.764	–0.117	–0.163 to –0.071	0.153	<.001
Illness intrusiveness	1–7	Lower	877	2.79	1.20	–0.082	–0.153 to –0.005	0.068	.03
Hypoglycemic symptoms	0–7	Lower	865	1.36	1.40	–0.188	–2.78 to –0.108	0.134	<.001
Fatigue	1–10	Lower	884	4.84	2.34	–0.508	–0.657 to –0.341	0.217	<.001
Sleep	1–10	Lower	882	3.93	2.95	–0.061	–0.240 to 0.119	0.021	.51
Aerobic exercise (min/week)		Higher	876	81.6	101	8.91	2.39 to 15.4	0.088	.01
Medication adherence	0–4	Lower	878	1.05	1.08	–0.135	–0.201 to –0.073	0.125	<.001
Communication with doctor	0–5	Higher	877	2.78	1.17	0.221	0.154 to 0.283	0.189	<.001
Proportion receiving recommended test in last 6 months[d]									
Eye exam	0,1	Higher	877	0.451	0.498	0.100	0.057 to 0.145	0.201	<.001
Foot exam	0,1	Higher	877	0.551	0.498	0.132	0.096 to 0.166	0.265	<.001
Cholesterol exam	0,1	Higher	877	0.763	0.425	0.090	0.061 to 0.127	0.222	<.001
Kidney exam	0,1	Higher	877	0.614	0.487	0.152	0.115 to 0.190	0.312	<.001

[a]HbA_{1c}: hemoglobin A_{1c}.

[b]Because of a laboratory recalibration changing HbA_{1c} measurement, HbA_{1c} prior to June 2014 was adjusted by adding 0.4.

[c]PHQ-8: 8-item Patient Health Questionnaire.

[d]"Don't know" set to "no examination."

Most health status and behavior variables showed statistically significant 6-month improvements. This remained when adjusting for multiple comparisons (using a *P* value of .003 as the criterion for significance).

For intent-to-treat analyses (not shown), we substituted a value of 0 (no change) for 6-month change scores for those who did not complete 6-month questionnaires. The intent-to-treat analyses of changes resulted in no differences from the *P* values shown in Table 2.

For all subsequent analyses, we included only those who completed 6-month questionnaires. We present these analyses as a means of determining who benefited in this heterogeneous population.

Increase in Receiving Recommended Tests (Hypothesis 4)

Of those who had not been examined (tested) in any of the 4 areas (eye, foot, cholesterol, kidney) in the year before baseline (n=85), only 12 (14%) had no tests in the next 6 months (not shown in tables). Thus, 73 (86%) of those who had not been examined at all had at least one examination in the 6 months after the program. A total of 46 (54%) of these had 3 or more of the 4 tests.

Differences in Changes by Type of Program or Location (Hypothesis 5)

There was little evidence of differences between participants in the two types of programs (Internet vs community). Only 1 change score differed by program: community participants had greater improvement in sleep (*P*<.001). Similarly, there was little evidence of differences by location of community programs. The only difference by community was for improvements in foot examinations, where community participants in St. Louis had marginally less improvement than in the other 2 communities (*P*=.04).

Analyses of Specific End Points (Hypothesis 6)

For each of the 5 outcomes described in hypothesis 6 above, we look at 6-month changes for those who exhibited symptoms and those who did not (not shown in the tables). Then we looked at those who benefited for 1 or more of these 5 problems.

HbA1c ≥9

Of those who supplied a blood sample for HbA_{1c} determination at both baseline and 6 months, 20.0% (98/489) had a baseline HbA_{1c} of ≥9. The proportion with HbA_{1c} >9 was reduced to 15.3% at 6 months (n=74, *P*<.001 from chi-square). The mean reduction in HbA_{1c} for the group starting at ≥9 was –0.93 (*P*<.001, effect size 0.73). For those with <9 at baseline, there was an increase of 0.05 in HbA_{1c} (*P*=.26).

Depression

At baseline, 22.0% (193/877) had symptoms indicating clinical depression (PHQ-8 score of ≥10) [24]. This was reduced to 16.3% at 6 months (n=144, $P<.001$). Baseline mean PHQ for the "depressed" group was 13.8 (SD 3.34). At 6 months, the mean reduction was 3.87 ($P<.001$, effect size 1.1); and 51.3% of those originally with a score ≥10 (n=99) had scores of <10. For those who had a score of <10 at baseline, there was a mean reduction of –0.21 in PHQ scores ($P=.08$)

Hypoglycemia

At baseline 332/865 (38.4%) had 2 or more symptoms of hypoglycemia (mean 2.88, SD 1.04). At 6 months, the number with 2 or more symptoms was reduced to 280 (32.4%). The 6-month change in mean hypoglycemic symptoms was –0.91 for those who had 2 or more symptoms at baseline ($P<.001$) and was an increase of 0.25 for those who had fewer than 2 symptoms at baseline ($P<.001$).

Medication-Taking Adherence

At baseline, 35.0% (n=307) were nonadherent in taking medicine (≥2 on a 0–4 scale), and at 6 months, 29.4% (n=260) were nonadherent ($P<.001$). For participants who were nonadherent at baseline, adherence improved by a mean of 0.73 (SD 1.05, $P<.001$, effect size 0.59) at 6 months, and 124 (40.4%) of the previous nonadherent participants were considered adherent. For those who had been adherent at baseline (n=575), there was an increase of 0.18 in mean nonadherence ($P<.001$).

Exercise

At baseline, 23.2% (203/876) indicated that they were taking no aerobic exercise. This was reduced to 18.4% (n=163) at 6 months ($P<.001$). At 6 months, the mean increase in exercise for nonexercisers was 43 minutes (SD 73, $P<.001$) of exercise per week, and 124 (61.1%) of the nonexercisers reported some aerobic exercise. For those reporting some exercise, there was a mean decrease of 1.4 minutes per week ($P=.61$).

In each of the cases above, the improvements for the less well-off group (high HbA_{1c}, indications of depression, etc) were much greater than any negative changes for those who had been doing well at baseline.

Percentage of Participants With at Least One Problem

At baseline, 70.4% (622/884) of participants fell into 1 or more of the above groups (high HbA_{1c}, high depression, hypoglycemic, nonadherent, or nonexercisers); 40.6% had 2 or more problem scores and 19.5% had 3 or more. The 5 criteria variables were not correlated with each other, with the exception of hypoglycemic symptoms and PHQ depression ($r=.37$).

Percentage of Participants Benefiting

We then looked at what proportion of the total study population completing 6-month questionnaires improved in at least one of the 5 criteria variables (hypothesis 7). Using effect-size improvements of at least 0.4 as an indication of an improvement, 75.0% of the study population (n=662) improved in at least one of the 5 criteria variables, and 37.1% (n=327) improved in 2 or more. The mean number of improvements (of ≥0.4 effect size) was 1.13. When broken down by how many of the 5 of these

problems a participant had, the mean number of improvements tended to increase with the mean number of problems. Those with none of the 5 criteria problems had a mean of 0.671 improvements (out of a possible 5, n=268); those with 1 problem or condition had a mean of 1.08 improvements (n=272); those with 2 problems had 1.31 improvements (n=195); those with 3 problems had 1.52 improvements (n=125); and those with 4 or 5 problems had 2.06 improvements (n=52). There was a correlation (Pearson $r=.35$) between the number of problems and the number of improvements ($P<.001$).

Discussion

There were modest but statistically significant improvement in 13 of the 14 outcome measures, and 10 of those were significant at the $P<.001$ level. There were significant increases in those completing suggested laboratory tests for diabetes. There was no significant change in quality of sleep. This consistency and level of significance suggests that these improvements were not the result of multiple testing. If we apply a Bonferroni correction and use .003 as the level of significance, 11 of 17 outcomes remain statistically significant.

BCBH-D may help meet the Healthy People 2020 diabetes objectives. The proportion of study participants with $HbA_{1c} >9$ was reduced from 20% (n=98) to 15% (n=74), and the proportion of participants tested for foot and microalbumin (kidney function) in the last 6 months increased by 13% and 15% (Table 2), respectively. This program also adds to the number receiving formal diabetes education.

Because this is a translation study, participants were not screened for HbA_{1c} or other symptoms. This led to a high degree of heterogeneity among study participants. In addition, many participants were already managing their diabetes and had little room for improvement. It was important, therefore, to determine who might benefit. It appears that those who had an HbA_{1c} of ≥9 decreased their HbA_{1c} by approximately the same amount as one would expect by taking metformin [30]. In addition, 75.0% (n=662) of the sample improved by an effect size of 0.4 or more for at least one of HbA_{1c}, depression, hypoglycemia, adherence to medications, and minutes of exercise. Those with none of the 5 problems still tended to benefit, with a mean improvement in 0.67 out of 5 possible criterial outcomes. However, as might be expected, those with more "problems" had more benefits (ie, making a greater mean number of improvements of ≥0.4 effect size among the 5 criterial measures).

6-Month Noncompleters

Considering the large initial sample (N=1242) and large number of outcomes, there were few statistical differences between 6-month completers and noncompleters. Based on the significant difference in baseline medication adherence, there is some evidence that noncompleters were slightly more likely to be medication noncompliers. They were also slightly younger. There is no evidence that they were more ill or had greater severity of symptoms, which would suggest that attrition would have been unlikely to bias 6-month outcomes.

Modes of Delivery

Despite a few differences at baseline, there were few differences in 6-month changes between those attending community workshops and the Internet group. This suggests that both modes are similarly effective. In this study, very few people had a choice of modes. However, it may be that future studies will find a greater population penetration when people are offered a choice.

Limitations

Because we lack a control group, we cannot be certain that the improvements observed in this study are not due to other factors. Plausible alternative explanations for the improvements include the introduction of new medications in the marketplace and health plan initiatives that were implemented for the purpose of improving the health of individuals with chronic conditions. It is also possible that there was a beneficial effect from participating in the workshops resulting in a greater likelihood of participants taking advantage of the Anthem initiatives and new medications. While these other factors are important to keep in mind, the consistency of the statistically significant improvements across multiple domains suggests a favorable impact of workshop participation.

By looking at the participants who were worse off on specific measures, we risk that part of the observed improvements might have been the result of regression to the mean. While regression to the mean may have contributed to subgroup improvements, it most likely does not explain all of the improvement.

Most of this study population were covered by a health plan and had a high mean baseline education (15.4 years). This may have contributed to the positive results. However, there were similar results when the small-group program was used with a low-education (mean of 7 years) Spanish-speaking population [8].

Data on how well improvements are sustained after 6 months are not included in this study. A future study will examine 12-month outcomes and address the sustainability of improvements following the intervention.

In order to reduce participant burden, unfortunately we did not include self-efficacy to manage diabetes in the questionnaires for this study. The program was designed to enhance self-efficacy and has been found to be significantly associated with self-efficacy in earlier randomized trials [7-9]. In a future translation study, it would be desirable to include a measure of self-efficacy to manage diabetes [31].

We did not attempt to control for changes in medication usage. Choosing to take medication for diabetes may be a part of self-management, and thus adding medication could be a positive outcome. In other cases, participants may have begun taking medication because of worsening health. A future study might want to look carefully at the relationships between self-management education, medication change, and health outcomes, as well as at other changes in behaviors as mediating variables.

Conclusion

As a community-based public health intervention, BCBH-D offered in two modes demonstrated small but significant benefits. It also showed promise for helping to meet at least some of the Healthy People 2020 diabetes objectives. More important, it demonstrated clinically significant benefits for those with high HbA$_{1c}$, and important benefits for those with depression and hypoglycemia, as well as nonadherers to medication and nonexercisers. The benefits differed by individual, but a large majority of the population demonstrated meaningful improvements in at least one of the above areas. This study demonstrated that the peer-facilitated BCBH-D in both face-to-face and Internet formats can assist a national sample of health plan members in improving their diabetes management.

Authors' Contributions

KL designed and coordinated the study, and wrote and edited the manuscript. PLR researched the data, and wrote and edited the manuscript. JG researched data and edited the manuscript. RMT researched data and edited the manuscript. KE contributed to discussion and edited the manuscript. DDL managed and coordinated the study and edited the manuscript.

Conflicts of Interest

KL and DDL have the potential to receive royalties if the program is further disseminated. Other authors have no potential conflicts of interest.

References

1. U.S. Department of HealthHuman Services. Healthy People 2020. Washington, DC: U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion; 2016 Jan 25. URL: http://www.healthypeople.gov/2020/topics-objectives/topic/diabetes [accessed 2016-01-26] [WebCite Cache ID 6epYzLSbH]

2. Li R, Shrestha SS, Lipman R, Burrows NR, Kolb LE, Rutledge S, Centers for Disease ControlPrevention (CDC). Diabetes self-management education and training among privately insured persons with newly diagnosed diabetes--United States, 2011-2012. MMWR Morb Mortal Wkly Rep 2014 Nov 21;63(46):1045-1049 [FREE Full text] [Medline: 25412060]

3. Haas L, Maryniuk M, Beck J, Cox CE, Duker P, Edwards L, 2012 Standards Revision Task Force. National standards for diabetes self-management education and support. Diabetes Care 2014 Jan;37 Suppl 1:S144-S153 [FREE Full text] [doi: 10.2337/dc14-S144] [Medline: 24357210]

4. Powers MA, Bardsley J, Cypress M, Duker P, Funnell MM, Fischl AH, et al. Diabetes self-management education and support in type 2 diabetes: a joint position statement of the American Diabetes Association, the American Association of Diabetes Educators, and the Academy of Nutrition and Dietetics. Diabetes Educ 2015 Aug;41(4):417-430. [doi: 10.1177/0145721715588904] [Medline: 26047627]

5. Administration for Community Living. Future Definition of Evidence-Based. 2015 Dec 31. URL: http://www.aoa.acl.gov/AoA_Programs/HPW/Title_IIID/Index.aspx [accessed 2016-01-26] [WebCite Cache ID 6epZCJ7lp]

6. Evidence-Based Leadership Council. What is "Evidence-Based"?. 2016. URL: http://www.eblcprograms.org/evidence-based/what-is-eb/ [accessed 2016-01-26] [WebCite Cache ID 6epZIg0Il]

7. Lorig K, Ritter PL, Laurent DD, Plant K, Green M, Jernigan VBB, et al. Online diabetes self-management program: a randomized study. Diabetes Care 2010 Jun;33(6):1275-1281 [FREE Full text] [doi: 10.2337/dc09-2153] [Medline: 20299481]

8. Lorig K, Ritter PL, Villa F, Piette JD. Spanish diabetes self-management with and without automated telephone reinforcement: two randomized trials. Diabetes Care 2008 Mar;31(3):408-414. [doi: 10.2337/dc07-1313] [Medline: 18096810]

9. Lorig K, Ritter PL, Villa FJ, Armas J. Community-based peer-led diabetes self-management: a randomized trial. Diabetes Educ 2009;35(4):641-651. [doi: 10.1177/0145721709335006] [Medline: 19407333]

10. Steinsbekk A, Rygg L, Lisulo M, Rise MB, Fretheim A. Group based diabetes self-management education compared to routine treatment for people with type 2 diabetes mellitus. A systematic review with meta-analysis. BMC Health Serv Res 2012;12:213 [FREE Full text] [doi: 10.1186/1472-6963-12-213] [Medline: 22824531]

11. Ramadas A, Quek KF, Chan CKY, Oldenburg B. Web-based interventions for the management of type 2 diabetes mellitus: a systematic review of recent evidence. Int J Med Inform 2011 Jun;80(6):389-405. [doi: 10.1016/j.ijmedinf.2011.02.002] [Medline: 21481632]

12. Heinrich E, Schaper N, de Vries N. Self-management interventions for type 2 diabetes: a systematic review. Eur Diabetes Nurs 2015 Feb 17;7(2):71-76. [doi: 10.1002/edn.160]

13. Pereira K, Phillips B, Johnson C, Vorderstrasse A. Internet delivered diabetes self-management education: a review. Diabetes Technol Ther 2015 Jan;17(1):55-63. [doi: 10.1089/dia.2014.0155] [Medline: 25238257]

14. Zhang X, Yang S, Sun K, Fisher EB, Sun X. How to achieve better effect of peer support among adults with type 2 diabetes: A meta-analysis of randomized clinical trials. Patient Educ Couns 2016 Feb;99(2):186-197. [doi: 10.1016/j.pec.2015.09.006] [Medline: 26463156]

15. Piatt GA, Orchard TJ, Emerson S, Simmons D, Songer TJ, Brooks MM, et al. Translating the chronic care model into the community: results from a randomized controlled trial of a multifaceted diabetes care intervention. Diabetes Care 2006 Apr;29(4):811-817. [Medline: 16567820]

16. Bojadzievski T, Gabbay RA. Patient-centered medical home and diabetes. Diabetes Care 2011 Apr;34(4):1047-1053 [FREE Full text] [doi: 10.2337/dc10-1671] [Medline: 21447667]

17. Bandura A. Self-Efficacy: The Exercise of Control. New York, NY: WH Freeman; 1997.

18. Holman H, Sobel D, Laurent D, Gonzalez V, Minor M, Lorig K. Living a Healthy Life With Chronic Conditions: Self-Management of Heart Disease, Arthritis, Diabetes, Asthma, Bronchitis, Emphysema & Others. Palo Alto, CA: Bull; 2012.

19. Stanford Patient Education Research Center. Program Fidelity Manual, Stanford Self-Management Programs 2012 update. 2012. URL: http://patienteducation.stanford.edu/licensing/FidelityManual2012.pdf [accessed 2016-01-26] [WebCite Cache ID 6epZOvqrO]

20. Lorig K, González VM, Laurent DD, Morgan L, Laris BA. Arthritis self-management program variations: three studies. Arthritis Care Res 1998 Dec;11(6):448-454 [FREE Full text] [Medline: 10030176]

21. Pillay J, Armstrong MJ, Butalia S, Donovan LE, Sigal RJ, Vandermeer B, et al. Behavioral programs for type 2 ciabetes mellitus: a systematic review and network meta-analysis. Ann Intern Med 2015 Dec 1;163(11):848-860. [doi: 10.7326/M15-1400] [Medline: 26414227]

22.   CoreMedica Laboratories. Dried Blood Spot (DBS). 2016. URL: https://www.coremedicalabs.com/blood-collection.html [accessed 2016-01-26] [WebCite Cache ID 6epZSghvN]

23.   Idler EL, Angel RJ. Self-rated health and mortality in the NHANES-I Epidemiologic Follow-up Study. Am J Public Health 1990 Apr;80(4):446-452. [Medline: 2316767]

24.   Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1-3):163-173. [doi: 10.1016/j.jad.2008.06.026] [Medline: 18752852]

25.   Devins GM, Mandin H, Hons RB, Burgess ED, Klassen J, Taub K, et al. Illness intrusiveness and quality of life in end-stage renal disease: comparison and stability across treatment modalities. Health Psychol 1990;9(2):117-142. [Medline: 2331973]

26.   Piette JD. Patient education via automated calls: a study of English and Spanish speakers with diabetes. Am J Prev Med 1999 Aug;17(2):138-141. [Medline: 10490057]

27.   Ritter PL, González VM, Laurent DD, Lorig KR. Measurement of pain using the visual numeric scale. J Rheumatol 2006 Mar;33(3):574-580. [Medline: 16511926]

28.   Lorig K. Outcome Measures for Health Education and other Health Care Interventions. Thousand Oaks, CA: Sage Publications; 1996.

29.   Morisky DE, Green LW, Levine DM. Concurrent and predictive validity of a self-reported measure of medication adherence. Med Care 1986 Jan;24(1):67-74. [Medline: 3945130]

30.   Bolen S, Feldman L, Vassy J, Wilson L, Yeh H, Marinopoulos S, et al. Systematic review: comparative effectiveness and safety of oral medications for type 2 diabetes mellitus. Ann Intern Med 2007 Sep 18;147(6):386-399. [Medline: 17638715]

31.   Ritter PL, Lorig K, Laurent DD. Characteristics of the Spanish- and English-language self-efficacy to manage diabetes scales. Diabetes Educ 2016 Apr;42(2):167-177. [doi: 10.1177/0145721716628648] [Medline: 26846336]

Abbreviations

BCBH-D: Better Choices, Better Health-Diabetes
HbA1c: hemoglobin A1c
PHQ-8: 8-item Patient Health Questionnaire

Original Paper

Being Human: A Qualitative Interview Study Exploring Why a Telehealth Intervention for Management of Chronic Conditions Had a Modest Effect

Alicia O'Cathain[1], PhD; Sarah J Drabble[1], PhD; Alexis Foster[1], MPH; Kimberley Horspool[1], MSc; Louisa Edwards[2], PhD; Clare Thomas[2], PhD; Chris Salisbury[2], MD

[1]Medical Care Research Unit, School of Health and Related Research, University of Sheffield, Sheffield, United Kingdom

[2]Centre for Academic Primary Care, School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom

Corresponding Author:
Alicia O'Cathain, PhD
Medical Care Research Unit
School of Health and Related Research
University of Sheffield
Regent Court
30 Regent Street
Sheffield, S1 4DA
United Kingdom
Phone: 44 114 222 0770
Fax: 44 114 222 0749
Email: a.ocathain@sheffield.ac.uk

Abstract

Background: Evidence of benefit for telehealth for chronic conditions is mixed. Two linked randomized controlled trials tested the Healthlines Service for 2 chronic conditions: depression and high risk of cardiovascular disease (CVD). This new telehealth service consisted of regular telephone calls from nonclinical, trained health advisers who followed standardized scripts generated by interactive software. Advisors facilitated self-management by supporting participants to use Web-based resources and helped to optimize medication, improve treatment adherence, and encourage healthier lifestyles. Participants were recruited from primary care. The trials identified moderate (for depression) or partial (for CVD risk) effectiveness of the Healthlines Service.

Objective: An embedded qualitative study was undertaken to help explain the results of the 2 trials by exploring mechanisms of action, context, and implementation of the intervention.

Methods: Qualitative interview study of 21 staff providing usual health care or involved in the intervention and 24 patients receiving the intervention.

Results: Interviewees described improved outcomes in some patients, which they attributed to the intervention, describing how components of the model on which the intervention was based helped to achieve benefits. Implementation of the intervention occurred largely as planned. However, contextual issues in patients' lives and some problems with implementation may have reduced the size of effect of the intervention. For depression, patients' lives and preferences affected engagement with the intervention: these largely working-age patients had busy and complex lives, which affected their ability to engage, and some patients preferred a therapist-based approach to the cognitive behavioral therapy on offer. For CVD risk, patients' motivations adversely affected the intervention whereby some patients joined the trial for general health improvement or from altruism, rather than motivation to make lifestyle changes to address their specific risk factors. Implementation was not optimal in the early part of the CVD risk trial owing to technical difficulties and the need to adapt the intervention for use in practice. For both conditions, enthusiastic and motivated staff offering continuity of intervention delivery tailored to individual patients' needs were identified as important for patient engagement with telehealth; this was not delivered consistently, particularly in the early stages of the trials. Finally, there was a lack of active engagement from primary care.

Conclusions: The conceptual model was supported and could be used to develop further telehealth interventions for chronic conditions. It may be possible to increase the effectiveness of this, and similar interventions, by attending to the human as well as the technical aspects of telehealth: offering it to patients actively wanting the intervention, ensuring continuity of delivery by enthusiastic and motivated staff, and encouraging active engagement from primary care staff.

XSL•FO
RenderX

Introduction

The increasing prevalence of chronic conditions presents a challenge to health systems internationally in terms of the ability to meet patients' health care needs. There is interest in the potential of technology to address this challenge by offering an alternative to face-to-face care between health care professionals and patients [1]. Telemedicine or telehealth delivers health care at a distance using information and communication technologies for diagnosis, treatment, and prevention of health problems [1]. These technologies can be supported by different types of clinical and nonclinical staff and thus expand health care provision and increase access to care. Policy makers worldwide have enrolled large numbers of patients in telehealth schemes [1-3] and are evaluating telehealth programs [1,4].

Despite the promotion of telehealth internationally, evidence of benefit is mixed [5-7]. A large review of the effectiveness of telehealth for chronic conditions concluded that the evidence base is weak and inconclusive due to publication bias, short-term outcome measurement, and a lack of focus on cost-effectiveness [5]. A review of reviews of telehealth concluded that telehealth could be effective for the management of some chronic conditions, but that evidence is mixed with a need for larger studies [6]. A more recent review of interactive telehealth concluded that telehealth was effective for some chronic conditions, specifically heart failure and diabetes, but that evidence was inconsistent for other conditions [7].

The lack of consistency of the evidence base on telehealth could reflect a lack of theoretical underpinning for many interventions or problems with the quality of their evaluation. It has been recommended that large, rigorous evaluations of any new interventions are undertaken [8]. Furthermore, process evaluations undertaken alongside trials of complex interventions such as telehealth may enable researchers to understand why interventions succeed or fail by exploring context, mechanisms of action, and implementation of the intervention [9]. Qualitative research can contribute to this [10].

Researchers have started to address the need for large, pragmatic trials of theory-based telehealth for chronic conditions. Two large, linked, randomized controlled trials (RCTs) of a telehealth intervention, known as the Healthlines Service, which followed up patients for a year, focused on depression and on risk factors for cardiovascular disease (CVD) [11]. The trial targeting depression identified a moderate clinical benefit [12], whereas the trial focusing on reducing risk factors for CVD identified a partial effect; that is, improvement in some individual risk factors but not overall CVD risk score [13]. An embedded qualitative study was undertaken with both these trials with the aim of explaining the results of the trials [11]. In this paper, we report this embedded qualitative study to explore why the trials showed modest effects only and then discuss the implications of this for future use and evaluation of this type of telehealth intervention.

Methods

The telehealth intervention is described in Textbox 1. The 2 RCTs are described in Textbox 2.

Textbox 1. The intervention

The content and delivery of the Healthlines Service was underpinned by a conceptual model called the TElehealth in CHronic disease (TECH) model [14]. This model was constructed by the research team based on an extensive review of quantitative and qualitative evidence, a qualitative interview study with patients and staff experienced with telehealth or with chronic conditions [15] and a postal survey of patients' levels of interest in different forms of telehealth [16]. The TECH model builds on the Chronic Care Model [17] and proposes that telehealth interventions are most likely to be effective and acceptable if they address (1) engagement of patients and health professionals; (2) effective chronic disease management (including self-management, optimization of treatment, care coordination); (3) partnership between providers; and (4) patient and health system context. This model was used to design a telehealth intervention for 2 exemplar chronic conditions: depression and raised CVD risk.

The Healthlines Service was based on regular telephone calls over a 12-month period from a nonclinical health information advisor (HIA). The content of the calls was guided by scripts within computer software designed for the study. The HIAs also used motivational interviewing skills to encourage behavior change and improve self-management. Patients were encouraged to identify goals and then offered links to information about quality-assessed resources on the Internet. Some aspects of the intervention were condition specific. For patients with depression, the intervention included book-based or computerized cognitive behavioral therapy and access to a Web-based mental health network (Big White Wall). For patients with raised CVD risk, the intervention included blood pressure self-monitoring using loaned blood pressure monitors with automated feedback via a Web portal and advice about diet, exercise, and smoking cessation. For both conditions, patients' use of medication was reviewed by the HIA. Problems with medication adherence were addressed, where patients were not being treated in accordance with national guidelines; a treatment recommendation was sent to their general practitioner and copied to the Web portal where the patient could view it. The intervention was designed to complement primary care delivered in general practice. The intervention was delivered by NHS Direct, which was a national telephone-based service at the time of the study. The staff members delivering the intervention were experienced HIAs who were given additional training to deliver the Healthlines Service.

Textbox 2. The trials

The Healthlines Service was tested in 2 linked, pragmatic randomized controlled trials comparing the intervention plus usual care versus usual care alone. Usual care for depression was attendance at general practice, including use of medication and possible referral to psychological services. Usual care for raised CVD risk was attendance at general practice where patients might receive blood pressure monitoring, medication, and lifestyle advice. The trials were undertaken with adults with depression or raised CVD risk recruited from 43 general practices in 3 areas of England. Both trials were powered to detect odds ratios of 1.7 with 80% power [11]. In total, 609 patients were recruited to the depression trial. The primary outcome was response to treatment measured using the Patient Health Questionnaire (PHQ-9) [18] and defined as a reduction ≥5 points and score <10 after 4 months. The treatment response was higher in the intervention group than in the control group (27% (68/255) vs 19% (50/270), odds ratio=1.7 (95% CI 1.1-2.5; P=.02). Twenty-five percent (78/307) of patients received little or none of the intervention [12], which is a similar rate to other pragmatic trials of telehealth for depression.

Overall, 641 people were recruited to the CVD risk trial. The primary outcome was response to treatment defined as maintenance or reduction in 10-year risk of CVD (measured by QRISK2 score [19]) after 12 months. Participants receiving the intervention had a modest response to treatment compared with those receiving usual care (50% (148/295) vs 43% (124/291), respectively; adjusted odds ratio 1.3; 95% CI 1.0-1.9). The intervention was associated with reductions in blood pressure (difference in mean systolic -2.7 mm Hg (95% CI -4.7 to -0.6)) and weight (-1.0 kg (95% CI -1.8 to -0.3)) but not in cholesterol or smoking status. Eight percent (26/325) of intervention participants received little or none of the intervention, and a third (103/325) received the full course of intended telephone encounters over the course of a year [13].

We undertook a qualitative interview study alongside the 2 trials to explain the results of the trials. We planned to interview 3 groups who could reflect on the intervention: primary care staff working in collaboration with the intervention who could offer perspectives on its feasibility and acceptability to primary care; staff from the organization delivering the intervention (NHS Direct) who could offer perspectives on feasibility; and patients who had experienced the intervention who could reflect on its acceptability. We chose to use the data collection method of interviews because they allow in-depth exploration of individuals' perceptions.

Sampling

For the first group (primary care staff), we planned to sample 6 general practices, selected to include practices with populations from varying levels of deprivation. We had to widen our original sample from 6 to 13 general practices because it proved difficult to recruit sufficient numbers of primary care staff from the original set of practices. Within the 13 practices, we sampled purposively to reflect the range of relevant professionals offering primary health care to participants within the intervention arm of the trial: general practitioners (GPs) and practice nurses or health care assistants.

For the second group (NHS Direct staff), we sampled staff purposively to include those delivering the intervention to participants (Health Information Advisors—HIAs), those offering technical expertise for the intervention, and those involved in team and strategic management.

For the third group (patients), we first sampled patients purposively from the intervention arm of the trials to ensure half of interviewees were in the depression trial and half in the CVD risk trial. Next, we used maximum variation sampling so that patients were interviewed who differed in terms of gender, age, and levels of depression or types of CVD risk factors. As a large proportion of patients using the intervention for depression used little or none of the intervention (25% in depression trial vs 8% in CVD risk trial), we also interviewed some patients who had withdrawn from the depression intervention.

Data Collection

For primary care staff, we wrote to GPs and practice nurses in participating practices asking for consent for an interview. We interviewed primary care staff at different stages of the trial period to obtain a mix of views at an early and later stage of the intervention delivery. Interviews took place face to face at their workplace or by telephone if this was more convenient.

For NHS Direct staff, we approached senior managers to identify relevant staff. We interviewed staff in July 2013, around 12 months after the first depression participant was randomized and 8 months after the first CVD risk participant was randomized. This allowed staff to reflect on both the early and later stages of intervention delivery.

For patients, we contacted those recruited to the intervention arm of the trials who had consented to participate in the interview study during the trial recruitment process. We interviewed these patients after at least 4 months (depression) or 6 months (CVD risk) of experiencing the intervention to allow us to obtain reflections on different stages of their care. This was after the primary outcome measure had been collected in the depression trial (4 months) and after the first collection of follow-up outcome data (at 6 months) in the CVD risk trial. Patients who had withdrawn from the intervention were interviewed within 5 months of recruitment. Face-to-face interviews with patients took place at their home or an alternative venue, depending on their preference.

SJD undertook most of the interviews, with support from AF and KH. We obtained written informed consent from all interviewees. Regardless of interviewee type, the focus of the interviews was on the intervention. We asked about its perceived utility, problems arising, and issues that enhanced or hindered its operation in practice. In addition, we asked about the components of the conceptual model underlying the intervention: engagement, promoting self-management, treatment optimization, care coordination, partnership, and context. Interviews lasted on average 45 minutes for staff, ranging from 16 to 88 minutes, and 58 minutes for patients, ranging from 21 to 124 minutes.

Analysis

Interviews were digitally recorded and transcribed verbatim. The framework approach was used to analyze the data [20]. We read some transcripts from each type of interviewee for familiarization (stage 1 of framework analysis). We constructed a thematic framework based on reading these transcripts and the functions of context, mechanisms of action, and implementation important to process evaluations [9] (stage 2 of framework analysis). As this qualitative study was embedded within RCTs, we supplemented this approach with a framework of the use of qualitative research with trials [10]. This permitted further exploration of themes concerning the trial, outcomes, and the health conditions under study. Subthemes of the theme "mechanisms of action" were informed by the components of the TElehealth in CHronic disease (TECH) model: engagement, promoting self-management, treatment optimization, care coordination, partnership, and context [14].

SJD coded all transcripts to the thematic framework, adding emerging subthemes throughout this process (stage 3 of framework analysis). SJD, AOC, AF, and KH then read the text within each subtheme, paying attention to which interviewees contributed to each subtheme. The final stage of the framework approach—"mapping and interpretation"—involved consideration of relationships between themes and subthemes. As recommended, the analysis was undertaken before any team member knowing the outcomes of the trials [21]. Findings of the qualitative study were discussed among the research team in September 2014 before the trial results were known. We suggested in our conclusions from this analysis that the intervention would be effective because there was evidence within our data that components of the conceptual model helped some patients in both trials, and the intervention was implemented largely as planned. We also suggested that the complexity of patients' lives and how the intervention was implemented appeared to diminish its impact. Paying attention to the balance of issues, we predicted a small-to-moderate benefit for each trial. In a second stage of analysis, after the trial outcomes were known in December 2014, we used the findings of this qualitative work to help explain the results of the trials. This involved focusing on the themes we considered to be most relevant to the research question of why this intervention had produced a modest effect, while taking care to acknowledge the uncertainties around our explanation.

The trials and qualitative study were approved by the National Research Ethics Service Committee South West—Frenchay (reference 12/SW/0009) and had the following trial registrations: ISRCTN14172341 (depression) and ISRCTN27508731 (CVD risk).

Results

Description of Participants

We undertook 45 interviews in total, with 21 staff and 24 patients.

Staff

We interviewed 6 GPs, 5 practice nurses, 1 health care assistant, and 1 practice-based research nurse (13 in total) from 13 of the general practices that had participated in the trials. We approached practice staff who had been involved to some extent in the trials, for example, GPs who had screened lists of potential trial participants before recruitment. From a total of 24 primary care staff approached for interview, 7 GPs and 4 practice nurses declined primarily because they did not feel they had anything to say about the intervention.

We interviewed 8 staff from NHS Direct. This included 4 HIAs who had delivered the intervention for varying lengths of time. Two had worked in the Healthlines Service from the beginning, 1 for a few months, and 1 had been in post for a month at the time of the interview. We also interviewed a strategic manager who had been involved in leading the intervention development, a technical manager who had helped to develop the intervention, a supervisor of the HIAs and a team manager from the wider organization who was not directly involved with the intervention but who managed the HIAs as part of larger team. This latter interview was undertaken to explore the wider organizational context in which the intervention was delivered.

Patients

We approached 16 depression and 20 CVD risk trial participants to obtain 12 interviews with each group. Patients declined to participate because they said they were not interested (n=6), were too busy (n=3), could not be contacted (n=1), or did not attend the arranged interview (n=2).

Interviewees participating in the depression trial were interviewed a median of 8 months after randomization, varying between 5 and 10 months. There were 7 women, they were all white, and mainly middle-aged (age range 30-66 years). This generally reflected the demographics of participants in the depression trial. According to the baseline PHQ-9 classifications, 1 interviewee had severe depression, 4 interviewees had moderately severe depression, and the remainder had moderate depression. Four interviewees had formally withdrawn from using the intervention at the time of the interview.

CVD risk interviewees were interviewed a median of 8 months after randomization, varying between 3 and 11 months. They were mainly men (n=9), all white, and all older (age range 62-75 years). This demographic mix was largely in line with the participants in the CVD risk trial. They had a mix of CVD risk factors at baseline: 2 smoked, 9 had a body mass index (BMI) ≥ 30, and 8 had systolic blood pressure ≥ 140. Eight were on blood pressure–lowering medication. The CVD risk score (QRISK2) was high for all interviewees (as that was an inclusion criteria for the trial), ranging between 21% and 58%; 3 had a score higher than 45%. None of the interviewees had formally withdrawn from using the intervention at the time of the interview.

Overview of Findings

The findings are presented using the framework of mechanisms of action, context, and implementation. We show that interviewees perceived that the intervention was useful for some patients and described aspects of the intervention that they valued. However, contextual issues and problems with implementation negatively affected the impact of the

intervention. Quotes are accompanied by labels showing the type of staff or characteristics of patients.

Mechanisms of Action

Perceptions That the Intervention Was Useful for Some Patients

Interviewees perceived that the intervention had improved the health of some patients. First, staff delivering the intervention described individual patients reporting improved mood and weight loss. They did not describe the characteristics of these patients, but, instead, described the characteristics of patients who they perceived were not being helped by the intervention (see the following section). Second, some of the patients interviewed reported improvements in health, which they associated with the intervention. Among the patients in the depression trial, 6 described benefits such as feeling more positive because they had been shown ways to cope, had learned to share problems with their family, felt listened to, or felt that someone cared about them:

> *what I needed was a way of dealing with the great sadnesses and a way of coming to terms with it, and I think I have got that from [pause], from The Healthlines Study"* [Dep 8, female, aged 66, with moderate depression at baseline]

Nine of the twelve CVD risk interviewees had a BMI of 30 or over at baseline. Three of these reported weight loss, which they attributed to the intervention. They were delighted with the amount of weight they had lost since joining the study and described other positive consequences, including reduced blood pressure, ability to walk more easily, and having more energy. Some CVD risk interviewees reported making lifestyle changes that could affect CVD risk factors, such as exercising more, eating more healthily, and reducing alcohol intake. Four of the eight interviewees with high blood pressure at baseline (systolic above 140 mm Hg) reported lowered blood pressure, and another reported reduced use of blood pressure medication related to the intervention. Improvements in blood pressure were attributed to weight loss or introduction of blood pressure medication:

Interviewer: You have got high blood pressure I am presuming?

> *CVD participant: Not anymore.*

Interviewer: Not anymore, good (laugh)

> *CVD participant: Mainly thanks to this system* [CVD risk 8, male, aged 70, with high blood pressure at baseline]

Aspects of the Intervention Valued by Staff and Patients

When asked about the different components of the intervention, interviewees tended to describe their value and how they helped to improve health. That is, there was support for the conceptual model on which the intervention was based. For example, both the HIAs delivering the intervention and patients receiving it described the necessity and value of different aspects of the intervention aimed at encouraging patient engagement. This included the technical support for patients, which helped them to use computer-based aspects of the intervention, the continuity of contact with the same HIA, which helped to build rapport

with patients, and enthusiastic and motivated HIAs who made the effort to tailor the intervention to patients' needs:

> *it has been good to build up some kind of relationship* [CVD risk 11, female, aged 49 years, overweight at baseline]

There was also support for the value of the self-management aspect of the intervention. Most of the patients we interviewed described how the intervention helped them to develop self-management skills through raising awareness of their health problems and educating them about ways of dealing with those problems. As 1 patient put it, the intervention was about

> *helping myself to help myself* [Dep 2, male, aged 60 years, with moderately severe depression at baseline.]

> *I think it makes people realize that there are things that you can do on a day to day basis […] to bring [their blood pressure] down, if they are checking it that regularly for a purpose. You know, I went out for a walk this morning and my blood pressure was really good today, and things like that. It makes it very obvious in black and white right in front of them that the days when they are doing things, and being a bit more well-behaved if you like, that it does make a difference.* [Practice Nurse 113]

> *and then it just gives them something to work on and I make it clear to them all that they have to do the hard work themselves if they want to reach their target. And 8 times out of 10 next time I speak to them they have done it or the first thing they say to me is "well I have been eating off a smaller plate" and it is really nice to hear that.* [Health Information Advisor 2]

Finally, there was some evidence of medication optimization occurring. Some patients receiving the depression intervention were on antidepressants, and some CVD risk participants were on blood pressure medication and statins. Interviewees reported that the intervention impacted medication taking through HIAs prompting patients to discuss medication with their GP or through letters directly from HIA s to GPs.

Context

Individual Context: Lack of Fit With Perceived Need

A key contextual issue, which may have impacted the effectiveness of the intervention, was patients' desire to improve their health. Patients with depression and primary care staff reported long waits for access to usual care services such as counseling and cognitive behavioral therapy. Patients with CVD risk factors wanted to improve their health, and some of those who wanted to improve their lifestyle perceived a lack of advice about how to do this, although the practice nurses we interviewed said they offered this service.

However, there were indications that some patients did not understand what the intervention entailed when they signed up to join the trial and, in fact, had no interest in what was on offer once they had started the intervention. Some patients in the depression trial described the intervention as too superficial, not giving access to a therapist, or the same as previous

treatments because they had already tried cognitive behavioral therapy.

Some patients in the CVD risk trial reported low motivation to change their lifestyle; they had been interested generally in improving their health without necessarily understanding that this would entail them making lifestyle changes or had joined the study for altruistic reasons in terms of helping others through participating in research:

> *and, I thought, well it will help me, but it might help somebody else, that is the reason I had a go* [CVD risk 7, male, aged 74 years, with high blood pressure at baseline]

Other patients in the CVD risk trial had no intention of addressing a key CVD risk factor that led to their eligibility for the trial. In particular, 2 of our CVD risk interviewees were smokers at baseline. Both these reported no success with smoking cessation because they did not want to stop smoking:

> *do not bother, I smoke* [CVD risk 1, male, aged 62 years, smoker with high blood pressure at baseline.]

Health Information Advisors noted that few patients had reported giving up or cutting down smoking and that this was a difficult lifestyle issue to have an impact on. The staff we interviewed believed that if intrinsic motivation to change was absent, then patients, particularly those in the CVD risk trial, would find it difficult to make the necessary lifestyle changes in the timeframe in which the intervention was offered.

Individual Context: Lack of Fit With Patients' Lives

Patients in the depression trial tended to be of working age, whereas those in the CVD risk trial tended to be retired. These younger patients with depression were described by HIAs as too busy due to childcare and employment to engage with key aspects of the intervention such as the telephone calls and homework for the cognitive behavioral therapy. The HIAs wondered whether lack of engagement was due to their depression and their busy lives. They felt that those who did complete the cognitive behavioral therapy course obtained benefit from it, and so they wanted the inclusion criteria for the trial to focus on those who were really committed to making changes and engaging with the intervention:

> *The depression ones, a large, it seems to be a lot, to me, younger people, a lot more women, not all but they are rushing around, they do not have time, they forget they have got appointments, and whether it is part of depression or not I do not know, but they do not often, they do not answer the phone* [Health Information Advisor 1]

Some interviewees from the depression trial described serious ongoing life events such as the threat of losing disability and unemployment benefits, physical illnesses, or coping with family members and friends who were very ill or depressed. These issues caused stress on top of the depression, making engagement with the intervention difficult. According to the HIAs in our study, life events preventing engagement with the intervention appeared to be less of an issue for CVD risk patients. Our interviewees with CVD risk factors did not offer the same description of complex lives as our interviewees with

depression. The level of complexity of patients' lives may have been related to age because the patients in the CVD risk trial were older and many were retired. Only 1 of the CVD risk interviewees was still in full-time paid employment, and this interviewee did report finding it difficult to fit the intervention into their life.

Research Context: A Randomized Controlled Trial

The intervention was offered in the context of an RCT. The intervention for depression was ready for use at the beginning of the trial and needed little or no adaptation during the trial. However, interviewees from NHS Direct discussed delays in starting the CVD risk intervention at the beginning of the trial due to a number of technical problems with the intervention. This resulted in some patients waiting for several months between randomization and receiving the intervention. As specified in the trial protocol, the primary outcome of the CVD risk trial was measured 12 months after randomization. This resulted in measurement of 12-month outcomes before some patients had completed the intervention, which may have reduced the measured effect of the intervention for CVD risk.

Implementation of the Intervention

When we asked the 3 groups of interviewees about different components of the intervention, they not only described the value of these components (see previous sections) but also described how they occurred in practice. With the exception of 3 issues (described in the following sections), their descriptions aligned with the planned implementation of the Healthlines Service.

Continuity of Enthusiastic and Motivated Health Information Advisors

Continuity of care—ensuring the same HIAs talked to the same patient throughout their care—was one of the ways in which the intervention delivered the patient engagement component of the TECH conceptual model. This appeared to be very important to some patients we interviewed and was compromised in the early months of implementation. Interviewees from NHS Direct described how, during the earlier months of the intervention, they tested out a model of using staff part-time in the Healthlines Service and part-time in the wider organization. This made it difficult for the same HIA to contact the same patients and also caused challenges for HIAs trying to learn to use a technically complex intervention. It was also compounded by large numbers of patients entering the CVD risk at the same time. This lack of continuity compromised the ability of HIAs to actively tailor the intervention to different patients. As the intervention progressed, NHS Direct changed the model of provision to a small dedicated team of staff who were enthusiastic about the intervention and felt motivated to help patients to improve their health. The HIAs we interviewed were part of this dedicated team and described how they placed emphasis on providing continuity of care and tailoring the intervention to the needs of individual patients. However, they also described how continuity of care could not be fully delivered even in the later stages because the small team sometimes struggled to cover sickness absence and holidays while still providing appointments that suited patients.

The variation in implementation was evident in patients' descriptions of their experiences. Some of the patients we interviewed appreciated the relationship they had built up with an HIA, feeling listened to and cared for. Others described HIAs as "going through the motions," rather than attempting to tailor the intervention:

> because the spiel was exactly the same [CVD risk 10, male, aged 71 years, with high blood pressure at baseline.]

This latter group struggled to engage with the intervention. Indeed, 3 of the interviewees in the depression trial who expressed concern about a protocoled approach had withdrawn from the intervention.

Modification of Intervention Delivery During the Trial

NHS Direct staff described how continuing technical difficulties had to be sorted out during the early weeks of using the intervention for CVD risk. Health Information Advisors explained how they had to learn to make the software work in the context of an ongoing conversation with a patient, modifying the flow of the scripts that were built into the intervention to reduce repetition for patients. They also described how they made notes about patients outside the computerized system to help them set and monitor plans for patients.

Collaboration With Primary Care

The primary care staff we interviewed had little to say about specific aspects of the intervention. Health Information Advisors and patients described how GPs responded to prompts to consult with patients or change medication but also described how they did not take an interest in patients' experiences of the intervention or proactively contact the Healthlines Service about individual patients. There was also some evidence that communication between primary care and HIAs did not always reach the level of partnership intended by the conceptual model, which could cause confusion for some patients. For example, GPs did not necessarily agree with advice from the intervention, which was based on national guidelines:

> there was this one particular patient who was constantly being, it was being suggested that he be reviewed by the GP. And the GP was reviewing him, but it was still the same, you know, it was a bit, you know, flogging a bit of a dead horse really, because she was, the GP was very happy with the blood pressure. Healthlines Study staff were saying, oh, no, no, no you need to go and see the GP [...] and of course the patient is the one caught in the middle [Practice Nurse 111]

Discussion

Principal Findings

The interviewees described improved outcomes in some patients receiving the intervention. They attributed these improvements to the intervention, describing how components of the conceptual model on which the intervention was based helped to achieve benefit. Aspects of the intervention addressing patient engagement, self-management, and medication optimization

were valued. Implementation of the intervention occurred largely as planned. However, problems related to context and implementation may have reduced the size of effect. For depression, the context of patients' lives was often complex, resulting in these working-age patients sometimes being unable to engage with the intervention. Some patients also wanted a more therapist-based approach rather than the cognitive behavioral therapy on offer. For CVD risk, contextual issues included some patients joining the trial in the hope of improving their health generally, or altruistically helping with research, rather than being motivated to make lifestyle changes to address their specific risk factors. In addition, implementation was not optimal in the early part of the CVD risk trial as technical difficulties with the intervention were addressed and staff delivering the intervention adapted it for use in practice. For both conditions, enthusiastic and motivated staff members offering continuity of intervention delivery tailored to individual patients' needs were identified as important for patient engagement with telehealth, but this was not delivered consistently, particularly in the early stages of the trials. Finally, there was a lack of active engagement with the intervention from GPs in primary care. Although some of these issues related to the technological aspects of the intervention, most of them are related to human issues—the complexity of patients' lives and the need for skillful human support to complement the technology.

Strengths and Limitations

One key strength of this qualitative study was the inclusion of interviews with a wide range of stakeholders: staff offering primary care to patients, managers, and frontline staff delivering the intervention and patients who had used the intervention and those who had withdrawn from it. This greatly improved our understanding of the trial results and provided support for the use of the TECH conceptual model to underpin these kinds of interventions. There were 4 limitations. First, we could have used nonparticipant observation in combination with the interviews, such as listening to telephone calls and observing HIAs in their daily work, which may have helped to further understand implementation of the intervention. Second, although we felt that we achieved data saturation at the data collection stage for most of the groups we targeted, this was not the case for participants in the CVD risk trial because of the range of risk factors they had. Third, inclusion of other groups may have helped to further understand the intervention; in particular, HIAs who had left the service and patients in the control arm of the trials. Finally, we completed our data collection before the end of the intervention. The organizational context in which the intervention was delivered changed toward the end of intervention delivery. NHS Direct ceased to operate toward the end of the trials, although the intervention continued to be offered by the same HIAs working for a different organization. During the change in the organization hosting the service, there was a pause in service delivery for some patients, and this might have affected their engagement with the intervention. However, we did not have data from those delivering or receiving the intervention during or after this change.

Comparison With Prior Work

Some patients in the trials did not engage with the intervention: 25% of patients in the depression trial and 8% of patients in the CVD risk trial used little or none of the intervention [12,13]. These rates were smaller than a trial of a Web-based program for reducing CVD risk where almost half of the intervention users had dropped out at 12 months [22]. Interestingly, the qualitative research undertaken alongside that trial recommended the addition of human interaction to motivate and engage patients. Our qualitative study identified that motivated staff could enhance patient engagement, and that engagement was also dependent on human factors related to the patients. It identified that patients wanted help with their health, but not necessarily the intervention on offer, or did not see the intervention as a priority in their complex or busy lives. This finding is similar to a systematic review of computerized cognitive behavioral therapy—a core component of our intervention for depression—which identified that a median of 56% participants completed a full course and that personal circumstance was more commonly cited as the cause for noncompletion than difficulties with the technology or social background [23]. We know that a large proportion (82%) of invited patients chose not to participate in our trials in the first place [24]. Among those actively declining participation (rather than not replying to the trial invite), common reasons given were that they were too busy or they were not interested in the research. It was also the case that some patients agreed to participate who did not want the intervention on offer or whose lives were too complex to make use of it. Although efforts were made to communicate to potential participants in advance about what the intervention entailed, it is possible that the nature of the intervention was not described clearly enough and was misunderstood, or that participants held expectations of the intervention that differed from their experience. These patients might have declined to take part in the trial if they had known more about the content of the intervention and the efforts required of them.

Researchers are beginning to test ways of increasing the acceptance of Internet-based mental health interventions using an informational video [25]. This type of video may also be useful when recruiting patients for trials of telehealth interventions to help them make informed choices about participation. For example, people who smoke and do not want to stop might decline to participate if they understand that a key focus on the intervention is to help them reduce this risk factor. This may reflect the real world more because, in practice, patients tend to access smoking cessation services if motivated to stop smoking. It is possible that a future trial with more emphasis on communicating the content of the intervention, and the efforts required by patients to obtain benefit, might result in larger effect sizes than seen here.

The importance of the human aspect of telehealth, in terms of who delivers the intervention and how, was evident from these interviews. This "personal context" of factors, related to the practitioners involved, in terms of their perceptions of the relevance of and interest in the intervention, their skills, and their motivation has been identified as a type of context affecting how interventions work [26]. In our study, this personal context

of motivated HIAs appeared to facilitate patient engagement with the intervention through both developing rapport and tailoring the intervention. These 2 issues have been identified as mechanisms of action of telehealth for chronic conditions [27]. Other researchers have also identified the importance of continuity of the person delivering telehealth and their level of motivation during delivery [23,28-30]. This has also been identified as important for the self-management of chronic conditions more generally. For example, a recent systematic review of interventions for the self-management of asthma identified the importance of actively engaging patients and having motivated professionals delivering interventions [31]. This focus on the importance of motivated humans delivering telehealth has not been identified consistently. For example, qualitative research alongside an RCT of an educational Web-based tool to prevent problems in young people whose parents had mental health problems identified technical problems as the key barrier [32].

The lack of proactive engagement with the intervention from primary care was perhaps not too surprising given earlier interviews with practice nurses and GPs before developing the intervention [15]. These health professionals were ambivalent and often skeptical about the contribution of telehealth to the care of chronic conditions. The conclusion of this earlier research was that there was work to be done in terms of helping primary care health professionals to understand the changes in roles and new ways of working necessary to facilitate the introduction and integration of telehealth innovations into their services. Our conclusions after delivery of the intervention were similar, in that there is a need to develop better strategies for primary care engagement with telehealth. This lack of primary care engagement with interventions aimed at chronic conditions is not specific to telehealth interventions [33].

Implications

When delivering this or similar interventions in the real world, service providers may wish to consider communicating the content of the intervention clearly to prospective users and the amount of time and effort required by them to obtain benefit. They may also wish to ensure the service is provided by motivated staff who can offer continuity of care and tailor the intervention to patients' needs. Given the lack of engagement from primary care, it may also be helpful for future interventions to try to develop better strategies for primary care engagement that also take into consideration the heavy workload in general practice in the United Kingdom currently. These actions may increase the effect of this or similar interventions in the future. There are also implications for the treatment of other chronic conditions. The conceptual model for the intervention was supported by this qualitative research, and so could be used to develop further interventions tailored for different conditions. These interventions would have to undergo rigorous evaluation in RCTs. Finally, there is a methodological implication for trialists. Due to technical problems and delays, some aspects of the intervention were not fully functional during the early months of the trials, particularly for CVD risk. The possible implications of this are that participants in the early stage of the trials may have received an underdeveloped intervention. Feasibility testing before a full evaluation is an important aspect

of the evaluation of complex interventions [34], although finding sufficient resources to do so within a fixed research timeline and budget when the problems cannot necessarily be anticipated in advance may be challenging.

Conclusions

This qualitative research helped to explain why the outcomes of 2 linked trials were modest. The conceptual model of the intervention was supported and could be used to develop further telehealth interventions for chronic conditions. It may be possible to increase the effectiveness of this, and similar interventions, by attending to the human and the technical aspects of telehealth: offering it to patients actively wanting the intervention, ensuring continuity of delivery by enthusiastic and motivated staff, and encouraging active engagement from primary care.

Conflicts of Interest

AOC is a member of the funding panel that commissioned this project.

References

1. World Heatlh Organization. Telemedicine. Opportunities and developments in Member States. Report on the second global survey on eHealth. In: Global Observatory for eHealth Series. Geneva: World Health Organization; 2010.

2. Darkins A, Ryan P, Kobb R, Foster L, Edmonson E, Wakefield B, et al. Care Coordination/Home Telehealth: the systematic implementation of health informatics, home telehealth, and disease management to support the care of veteran patients with chronic conditions. Telemed J E Health 2008 Dec;14(10):1118-1126. [doi: 10.1089/tmj.2008.0021] [Medline: 19119835]

3. Cruickshank J. 2020health. 2012. Telehealth: what can the NHS learn from experience at the US Veterans Health Administration? URL: http://www.2020health.org/dms/2020health/downloads/reports/Telehealth-VA/Telehealth%20VA.pdf [accessed 2016-06-07] [WebCite Cache ID 6h3QIAq0p]

4. European Commission. Renewing Health: REgioNs of Europe WorkINg toGether for HEALTH. 2013. 9 Regions joining forces for revolutionising chronic patient care thanks to telemedicine URL: http://www.renewinghealth.eu/documents/28946/5c71382c-effb-4a6b-9eaf-5b3e98219a04 [accessed 2016-06-07] [WebCite Cache ID 6h3QjAuDk]

5. Wootton R. Twenty years of telemedicine in chronic disease management--an evidence synthesis. J Telemed Telecare 2012 Jun;18(4):211-220 [FREE Full text] [doi: 10.1258/jtt.2012.120219] [Medline: 22674020]

6. Ekeland AG, Bowes A, Flottorp S. Effectiveness of telemedicine: a systematic review of reviews. Int J Med Inform 2010 Nov;79(11):736-771. [doi: 10.1016/j.ijmedinf.2010.08.006] [Medline: 20884286]

7. Flodgren G, Rachas A, Farmer AJ, Inzitari M, Shepperd S. Interactive telemedicine: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2015;9:CD002098. [doi: 10.1002/14651858.CD002098.pub2] [Medline: 26343551]

8. Ekeland AG, Bowes A, Flottorp S. Methodologies for assessing telemedicine: a systematic review of reviews. Int J Med Inform 2012 Jan;81(1):1-11. [doi: 10.1016/j.ijmedinf.2011.10.009] [Medline: 22104370]

9. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. BMJ 2015;350:h1258 [FREE Full text] [Medline: 25791983]

10. O'Cathain A, Thomas KJ, Drabble SJ, Rudolph A, Hewison J. What can qualitative research do for randomised controlled trials? A systematic mapping review. BMJ Open 2013 Jun;3(6):e002889 [FREE Full text] [doi: 10.1136/bmjopen-2013-002889] [Medline: 23794542]

11. Thomas CL, Man M, O'Cathain A, Hollinghurst S, Large S, Edwards L, et al. Effectiveness and cost-effectiveness of a telehealth intervention to support the management of long-term conditions: study protocol for two linked randomized controlled trials. Trials 2014;15:36 [FREE Full text] [doi: 10.1186/1745-6215-15-36] [Medline: 24460845]

12. Salisbury C, O'Cathain A, Edwards L, Thomas C, Gaunt D, Hollinghurst S, et al. Effectiveness of an integrated telehealth service for patients with depression: a pragmatic randomised controlled trial of a complex intervention. Lancet Psychiatry 2016 Jun;3(6):515-525 [FREE Full text] [doi: 10.1016/S2215-0366(16)00083-3] [Medline: 27132075]

13. Salisbury C, O'Cathain A, Thomas C, Edwards L, Gaunt D, Dixon P, et al. Telehealth for patients at high risk of cardiovascular disease: pragmatic randomised controlled trial. BMJ 2016;353:i2647 [FREE Full text] [Medline: 27252245]

14. Salisbury C, Thomas C, O'Cathain A, Rogers A, Pope C, Yardley L, et al. TElehealth in CHronic disease: mixed-methods study to develop the TECH conceptual model for intervention design and evaluation. BMJ Open 2015 Feb;5(2):e006448 [FREE Full text] [doi: 10.1136/bmjopen-2014-006448] [Medline: 25659890]

15. Segar J, Rogers A, Salisbury C, Thomas C. Roles and identities in transition: boundaries of work and inter-professional relationships at the interface between telehealth and primary care. Health Soc Care Community 2013 Nov;21(6):606-613. [doi: 10.1111/hsc.12047] [Medline: 23656381]

16. Edwards L, Thomas C, Gregory A, Yardley L, O'Cathain A, Montgomery AA, et al. Are people with chronic diseases interested in using telehealth? A cross-sectional postal survey. J Med Internet Res 2014;16(5):e123 [FREE Full text] [doi: 10.2196/jmir.3257] [Medline: 24811914]

17. Wagner EH, Austin BT, Von KM. Improving outcomes in chronic illness. Manag Care Q 1996;4(2):12-25. [Medline: 10157259]

18. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [Medline: 11556941]

19. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Minhas R, Sheikh A, et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. BMJ 2008 Jun 28;336(7659):1475-1482 [FREE Full text] [doi: 10.1136/bmj.39609.449676.25] [Medline: 18573856]

20. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess RG, editors. Analyzing qualitative data. London: Routledge; 1994:173-194.

21. Oakley A, Strange V, Bonell C, Allen E, Stephenson J. Process evaluation in randomised controlled trials of complex interventions. BMJ 2006 Feb 18;332(7538):413-416 [FREE Full text] [doi: 10.1136/bmj.332.7538.413] [Medline: 16484270]

22. Watson S, Woodside JV, Ware LJ, Hunter SJ, McGrath A, Cardwell CR, et al. Effect of a Web-Based Behavior Change Program on Weight Loss and Cardiovascular Risk Factors in Overweight and Obese Adults at High Risk of Developing Cardiovascular Disease: Randomized Controlled Trial. J Med Internet Res 2015;17(7):e177 [FREE Full text] [doi: 10.2196/jmir.3828] [Medline: 26183659]

23. Waller R, Gilbody S. Barriers to the uptake of computerized cognitive behavioural therapy: a systematic review of the quantitative and qualitative evidence. Psychol Med 2009 May;39(5):705-712. [doi: 10.1017/S0033291708004224] [Medline: 18812006]

24. Foster A, Horspool KA, Edwards L, Thomas CL, Salisbury C, Montgomery AA, et al. Who does not participate in telehealth trials and why? A cross-sectional survey. Trials 2015;16:258 [FREE Full text] [doi: 10.1186/s13063-015-0773-3] [Medline: 26044763]

25. Ebert DD, Berking M, Cuijpers P, Lehr D, Pörtner M, Baumeister H. Increasing the acceptance of internet-based mental health interventions in primary care patients with depressive symptoms. A randomized controlled trial. J Affect Disord 2015 May 1;176:9-17. [doi: 10.1016/j.jad.2015.01.056] [Medline: 25682378]

26. Wells M, Williams B, Treweek S, Coyle J, Taylor J. Intervention description is not enough: evidence from an in-depth multiple case study on the untold role and impact of context in randomised controlled trials of seven complex interventions. Trials 2012;13:95 [FREE Full text] [doi: 10.1186/1745-6215-13-95] [Medline: 22742939]

27. Vassilev I, Rowsell A, Pope C, Kennedy A, O'Cathain A, Salisbury C, et al. Assessing the implementability of telehealth interventions for self-management support: a realist review. Implement Sci 2015;10:59 [FREE Full text] [doi: 10.1186/s13012-015-0238-9] [Medline: 25906822]

28. Greenhalgh T, Procter R, Wherton J, Sugarhood P, Hinder S, Rouncefield M. What is quality in assisted living technology? The ARCHIE framework for effective telehealth and telecare services. BMC Med 2015;13:91 [FREE Full text] [doi: 10.1186/s12916-015-0279-6] [Medline: 25902803]

29. Gambling T, Long AF. The realisation of patient-centred care during a 3-year proactive telephone counselling self-care intervention for diabetes. Patient Educ Couns 2010 Aug;80(2):219-226. [doi: 10.1016/j.pec.2009.11.007] [Medline: 20006458]

30. Knowles SE, Toms G, Sanders C, Bee P, Lovell K, Rennick-Egglestone S, et al. Qualitative meta-synthesis of user experience of computerised therapy for depression and anxiety. PLoS One 2014;9(1):e84323 [FREE Full text] [doi: 10.1371/journal.pone.0084323] [Medline: 24465404]

31. Pinnock H, Epiphaniou E, Pearce G, Parke H, Greenhalgh T, Sheikh A, Taylor Stephanie J C. Implementing supported self-management for asthma: a systematic review and suggested hierarchy of evidence of implementation studies. BMC Med 2015;13:127 [FREE Full text] [doi: 10.1186/s12916-015-0361-0] [Medline: 26032941]

32. Woolderink M, Bindels Jill A P M, Evers Silvia M A A, Paulus Aggie T G, van Asselt Antoinette D I, van Schayck Onno C P. An Online Health Prevention Intervention for Youth with Addicted or Mentally Ill Parents: Experiences and Perspectives of Participants and Providers from a Randomized Controlled Trial. J Med Internet Res 2015;17(12):e274 [FREE Full text] [doi: 10.2196/jmir.4817] [Medline: 26633244]

33. Coupe N, Anderson E, Gask L, Sykes P, Richards DA, Chew-Graham C. Facilitating professional liaison in collaborative care for depression in UK primary care; a qualitative study utilising normalisation process theory. BMC Fam Pract 2014;15:78 [FREE Full text] [doi: 10.1186/1471-2296-15-78] [Medline: 24885746]

34. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008;337:a1655 [FREE Full text] [Medline: 18824488]

XSL•FO
RenderX

Abbreviations

BMI: body mass index
CVD: cardiovascular disease
GP: general practitioner
HIA: health information advisor
PHQ-9: Patient Health Questionnaire
RCT: randomized controlled trial
TECH: TElehealth in CHronic disease

XSL•FO
RenderX

Original Paper

Participation in an Intensive Longitudinal Study with Weekly Web Surveys Over 2.5 Years

Jennifer Barber[1], PhD; Yasamin Kusunoki[1], PhD; Heather Gatny[2], MSc; Paul Schulz[2], MS

[1]Institute for Social Research, Population Studies Center, University of Michigan, Ann Arbor, MI, United States
[2]Institute for Social Research, Survey Research Center, University of Michigan, Ann Arbor, MI, United States

Corresponding Author:
Heather Gatny, MSc
Institute for Social Research
Survey Research Center
University of Michigan
426 Thompson St
Ann Arbor, MI, 48106
United States
Phone: 1 7346153560
Fax: 1 7346153560
Email: hgatny@umich.edu

Abstract

Background: Technological advances have made it easier for researchers to collect more frequent longitudinal data from survey respondents via personal computers, smartphones, and other mobile devices. Although technology has led to an increase in data-intensive longitudinal studies, little is known about attrition from such studies or the differences between respondents who complete frequently administered surveys in a timely manner, and respondents who do not.

Objective: We examined respondent characteristics and behaviors associated with continued and on-time participation in a population-based intensive longitudinal study, using weekly web-based survey interviews over an extended period.

Methods: We analyzed data from the Relationship Dynamics and Social Life study, an intensive longitudinal study that collected weekly web-based survey interviews for 2.5 years from 1003 18- and 19-year-olds to investigate factors shaping the dynamics of their sexual behavior, contraceptive use, and pregnancies.

Results: Ordinary least squares and logistic regression analyses showed background respondent characteristics measured at baseline were associated with the number of days respondents remained enrolled in the study, the number of interviews they completed, and the odds that they were late completing interviews. In addition, we found that changes in pregnancy-related behaviors reported in the weekly interviews were associated with late completion of interviews. Specifically, after controlling for sociodemographic, personality, contact information, and prior experience variables, we found that weekly reports such as starting to have sex (odds ratio [OR] 1.17, 95% CI 1.03-1.32, P=.01), getting a new partner (OR 1.76, 95% CI 1.53-2.03, P<.001), stopping the use of contraception (OR 1.28, 95% CI 1.10-1.49, P=.001), and having a new pregnancy (OR 5.57, 95% CI 4.26-7.29, P<.001) were significantly associated with late survey completion. However, young women who reported changes in pregnancy-related behaviors also had lower levels of study attrition, and completed more interviews overall, than did their counterparts.

Conclusions: We found that measures of participation in a longitudinal study with weekly web surveys varied not only by respondent characteristics, but also by behaviors measured across the surveys. Our analyses suggest that respondents who experience the behaviors measured by the study may maintain higher participation levels than respondents who do not experience those behaviors.

XSL•FO
RenderX

Introduction

Intensive longitudinal data are measures collected from participants across multiple observations. The number of observations that make data *intensive* is subject to debate, and "may be in the tens, hundreds, or thousands" [1]. The first intensive longitudinal data collections used pagers (beepers) to collect timed or event-driven data [2]. These studies were based on small, non-representative samples (eg, 75 high school students). Today, the widespread use of personal computers, smartphones, and other mobile devices greatly facilitate intensive data collection on larger and more diverse samples [1,3,4]. As these technologies become less expensive and more pervasive, intensive longitudinal data collection is likely to become even more common in survey research [5].

One obvious risk for studies with frequent measurement is the potential that some respondents will feel burdened by the task and will not continue in the study, or will respond sporadically to measurement attempts. If these respondents are different from those who remain in the study or who participate more consistently, then bias becomes a threat to the study [6]. Much research has addressed the causes and consequences of continued participation in multi-wave studies, as attrition affects the extent to which the sample remaining at the end of the study represents the initial sampling frame [7-14]. Little research, however, has addressed attrition in a population-based *intensive* longitudinal study with weekly interviews over a multi-year period.

Analyses of respondent characteristics and behaviors affecting on-time participation, an issue that is especially crucial to longitudinal designs with frequent measurement, are also lacking. Computer-based interviewing methods may tailor survey questions to reflect late or skipped participation by referring to the period since the respondent's last interview. However, at some point, late participation in frequent data collections weakens the data set. For example, a weekly survey in which each survey is completed four weeks late becomes a monthly survey. Little is known about the differences between respondents who complete frequent surveys in a timely manner and respondents who do not.

In this analysis, we examined factors associated with continued and on-time participation in the Relationship Dynamics and Social Life (RDSL) study. The RDSL study collected *weekly* data from 18- and 19-year-old women on their attitudes, relationships, sexual behaviors, contraceptive use, and pregnancy status for 2.5 years, amounting to 130 observations from each fully participating respondent by the end of the study. By asking respondents to recall events over just the prior week, the study greatly increased researchers' ability to determine the sequences of events leading to pregnancy, and to identify reciprocal relationships among attitudes and behaviors.

Using the RDSL data, we analyzed variance by individual characteristics and behaviors in the number of days respondents remained enrolled in the study, the number of survey interviews they completed, and the odds that they were late completing survey interviews. The individual variables were obtained from respondents' self-reported characteristics at baseline, and

subsequent pregnancy-related behaviors and experiences were reported during the course of the study.

Methods

The Relationship Dynamics and Social Life Study

The RDSL used a population-based sample of 1003 respondents randomly selected from a list of 18- and 19-year-old women registered as having either a driver's license or a Personal Identification Card (PID) in a single county in Michigan, USA. The sampling method was a cost-based decision; due to the relatively sparse distribution of 18- to 19-year-olds in the general population, random selection from this pool allowed us to avoid screening a large number of households. Comparison of the driver's license and PID data by zip code to 2000 census-based population projections revealed 96% agreement between the frame count and the projections for this population (authors' calculations).

A 60-minute face-to-face baseline survey interview was conducted between March 2008 and July 2009 to assess important aspects of family background, demographics, attitudes, romantic relationships, education, and career trajectories. At the conclusion of this baseline interview, respondents were invited to participate in a *journal* that was a structured survey interview every week for 2.5 years, focused on measures of pregnancy desires, contraceptive use, pregnancy, and relationship characteristics/behaviors such as commitment and sex. Respondents could elect to complete journal interviews on the web or with an interviewer via phone. Ninety-two percent of women who agreed to complete the journal had internet access and were encouraged to complete their journal surveys by web. The journal portion of the study concluded in January 2012.

The response rate for the baseline survey was 83.72% (1003/1198) [15]. Almost all of those who completed the baseline survey enrolled in the journal portion of the study (98.9%, 992/1003). The first journal was completed with the help of the professional interviewer immediately following the baseline interview. Of the 992 respondents who enrolled, 953 (96.1%) completed a second journal on their own. Of these 953, 741 (77.8%) participated in the journal interviews for at least 18 months and 604 (63.4%) completed their final interview at 900 or more days after enrollment (2.47 years). Figure 1 shows the percentage of respondents who remained enrolled in the study from 30 days after enrollment to 900 or more days after enrollment. Of the 128,960 possible weekly journals the study aimed to collect across 2.5 years from the 992 enrolled respondents, 58,594 (45.44%) were completed. Considering weekly survey questions were adjusted to refer to the period since the last interview for up to two weeks between surveys, this resulted in only a modest number of missing weeks. Item-specific missing data in the weekly interviews was also quite low at 3%.

RDSL researchers took several steps to minimize attrition and nonresponse in the study. First, we provided three types of participant incentives: money, tokens of appreciation, and regular reports on our study findings. Respondents were paid

$1 per weekly journal with $5 bonuses for on-time completion of five weekly journals in a row. Journal incentives were distributed via reloadable, prepaid debit cards. Our prior research suggests that the ease and speed of the journal incentive payments encouraged continued participation [16]. Respondents also received small tokens of appreciation for their continued participation in the journal study (eg, a pen, compact, lip balm). We also provided regular reports to participants on current study findings, in large part because our previous ethnographic work with 18-year-olds indicated their desire to be kept apprised on the research to which they were contributing.

Second, we developed a multi-modal system of contacts to motivate and remind respondents to complete their weekly journals. The system, summarized in Figure 2, includes manual and automatic participant contact points that are recurrent and varied by mode, to increase response rates [17-20]. The day a journal is completed is day 0, with the next journal made available on the study's website five days later, on day 5. On day 7, respondents were sent an automated invitation to complete the next journal via their selected contact mode (text messaging, email, or both). These automated invitations recurred on days 8 and 9, and on day 10 the reminder mode was switched to a telephone call. A second reminder call was made on day 12 and, if the journal was still not completed, a new automated invitation was sent via email or text message on day 14. This pattern of invitations and reminders was repeated until the respondent completed another journal interview or explicitly asked to be removed from the study. For participants not responding by day 30, we supplemented the automated email and text message reminders by mailing *refusal-conversion packets* to their homes that included a letter and a small gift (eg, pen, compact, lip balm). In addition, we offered respondents cash bonuses by phone/email for completing the next journal of $10 at day 60, $20 at day 90, and $30 at day 120.

Finally, to decrease respondent burden, we kept the survey short (approximately five minutes) and included only questions that were essential to the study (determinants of pregnancy). For instance, we did not include questions on finances, which respondents tend not to want to answer [11,21]. See Barber et al. (2011) for more information regarding the design and implementation of the RDSL study [22]. Sample journal interviews are provided in Multimedia Appendix 1.

Figure 1. Continued participation in the journal.
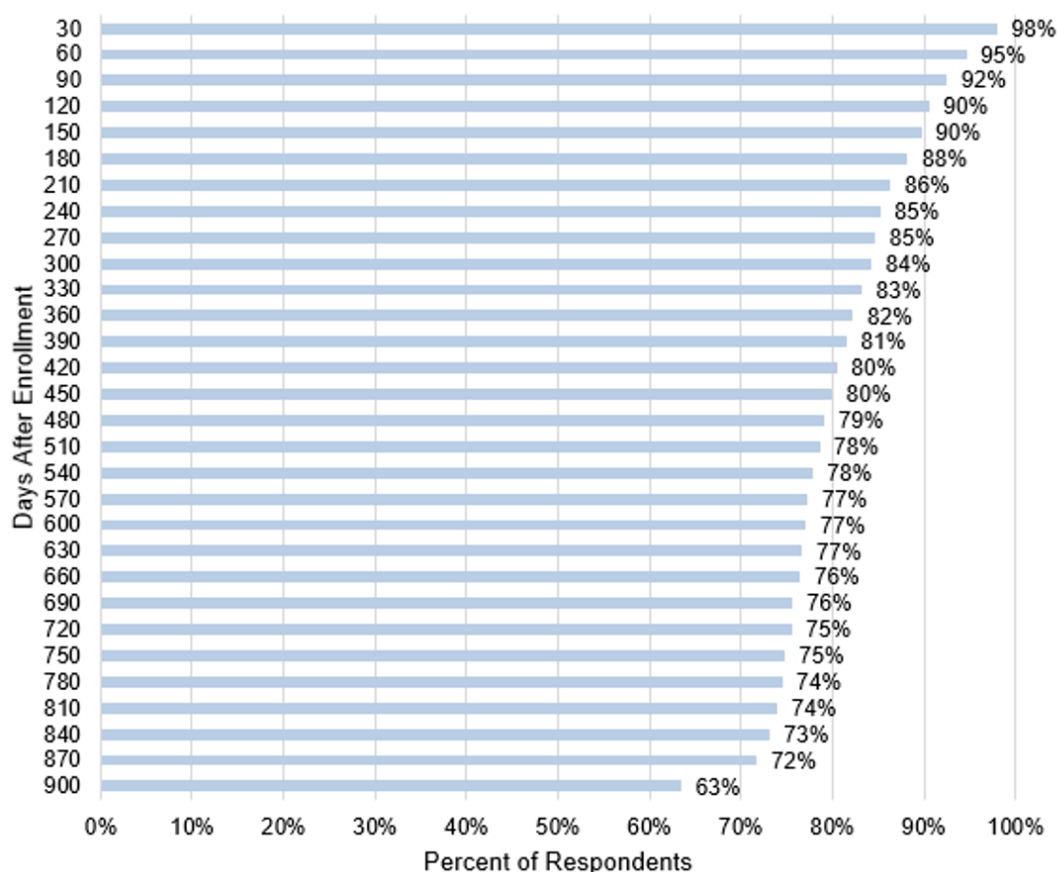
Figure 2. Heuristic of reminder protocol.
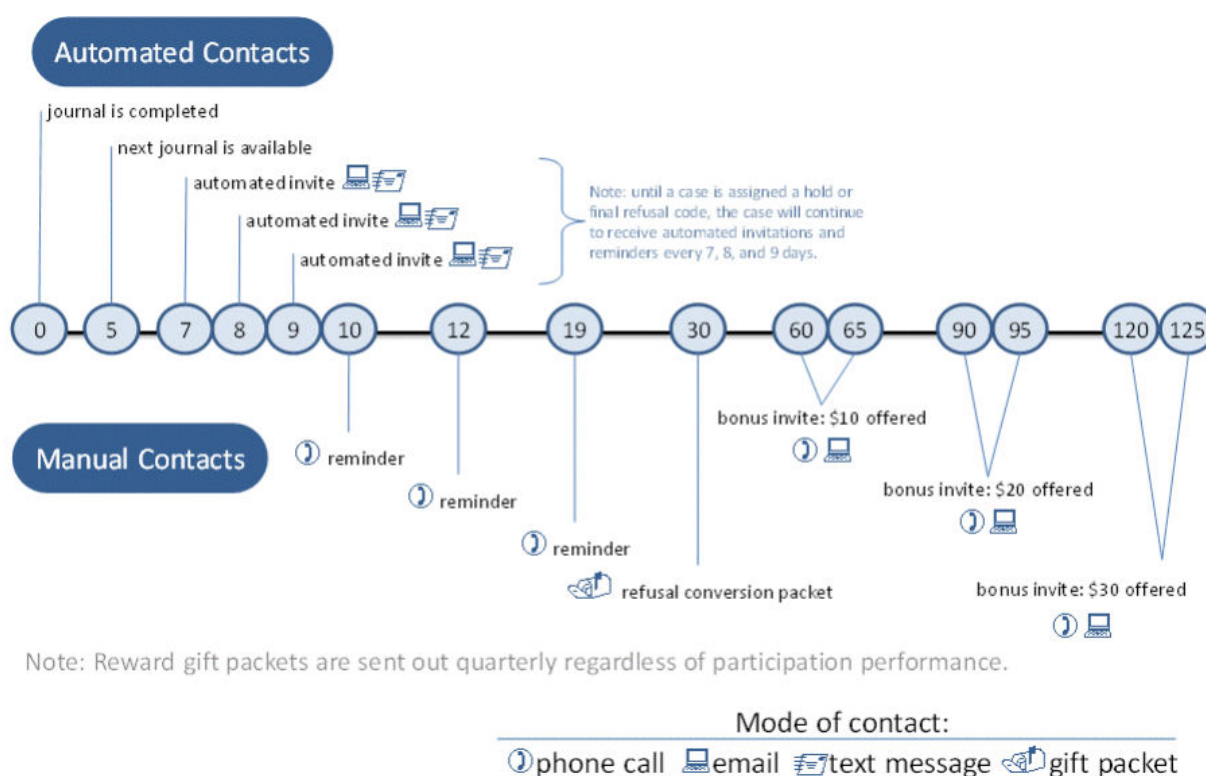
Measures

Non-Time-Varying Individual Characteristics

For the 953 RDSL study respondents who completed at least two journal interviews, Multimedia Appendix 2 shows mean measures for five sets of individual-level variables collected during the baseline survey and journal: (1) sociodemographic characteristics, (2) personality characteristics, (3) contact information/mode, (4) adolescent experiences (prior to the study) related to pregnancy, and (5) changes summarized over the study period. We explored the relationship of these non-dynamic variables to weekly survey participation.

Sociodemographic Characteristics

Previous research has documented that respondents from low socioeconomic backgrounds, and minority respondents, are more likely to drop out of panel studies [23,24]. In these analyses, we investigated a large set of sociodemographic characteristics we have included in other studies [25-31]. As shown in Multimedia Appendix 2, approximately one-third of the analytic sample was African American. The very few Latinas in the sample were coded according to their answer to the question of their race (some selected African American, while others selected white). To measure education of these 18- and 19-year-olds, we used a categorical variable combining enrollment and attainment: 13.5% (129/953) were enrolled in high school, 28.5% (272/953) were enrolled in a 2-year college/vocational program, 27.7% (264/953) were enrolled in 4-year college, 21.9% (209/953) had completed high school but were not enrolled in further education, and 8.3% (79/953) dropped out of high school and were not currently attending school. More than one-quarter (26.3%, 251/953) of the

respondents were receiving public assistance. On average, respondents rated the importance of religion in their lives as 2.69 on a scale of 1 (not important) to 4 (more important than anything else). Over one-half (52.4%, 499/953) grew up with two parents, 39.8% (379/953) with one biological parent only (no step-parent), and 7.9% (75/953) in another arrangement (eg, with grandparents or an aunt). Respondents were relatively equally distributed across the four parental income categories, with the highest proportion (27.9%, 266/953) in the $15,000 - $44,999 category and 20.1% (192/953) not knowing their parents' income. The average age at baseline was 19.19 years, with a median age of 19.

Personality Characteristics

Previous research has also documented the relationship between personality characteristics and study attrition. Attrition tends to be low among respondents who are assessed to be agreeable and/or conscientious, and higher among respondents who are extraverted [32]. The baseline interview included a series of questions adapted from the Neuroticism, Extraversion, and Openness to Experience-Five Factor Inventory (NEO-FFI), which measures extraversion, agreeableness, conscientiousness, neuroticism, and intellect/imagination [32]. Of the NEO-FFI's 60 questions regarding personality traits, we included only the 40 used in Add Health Wave IV, presenting them as statements with five possible answers coded from 1 (strongly disagree) to 5 (strongly agree). As shown in Multimedia Appendix 2, respondents were distributed fairly evenly across the five personality categories using this coding scheme, with agreeableness and neuroticism ranking the highest and lowest in terms of respondents' identification with the characteristic.

Contact Information/Mode

We also examined whether contact information provided, mode of the reminder, and mode of the journal interview were related to survey participation. Respondents were asked during the first journal interview (completed with the professional interviewer immediately following the baseline interview) to provide a home phone number, cell phone number, and an email address, with 84.1% (801/953) providing both an email address and a phone number and 15.9% (152/953) providing only one or neither. Respondents also selected their preferred automated reminder method for future journals, with 33.2% (316/953) electing to receive both a text and an email, and 66.8% (637/953) choosing one or neither reminder method. In addition, each week of the survey, respondents could elect to complete their interview via phone or web, with respondents ultimately completing 11.77% (6778/57,602) of all interviews by phone. Many of these phone call completions resulted from an interviewer-initiated call made on day 10 (the first day of manual contacts) in an attempt to complete the interview by phone. Thus, interviews completed by phone had a higher fraction late (1341/6778, 19.78%) than those completed on the web (5035/50,824, 9.91%).

Adolescent Experiences Related to Pregnancy

Although previous research has not linked pregnancy-related experiences to study attrition, we explored several salient experiences that occurred *prior to the study period*. At the baseline interview, 51.5% (491/953) of respondents reported having had sex by age 16, 59.7% (569/953) reported two or more sexual partners prior to the study period, 48.1% (458/953) reported having had sex without contraception, and 26.0% (248/953) reported having a prior pregnancy (17.0% [162/953] had one prior pregnancy and 9.0% [86/953] had two or more).

Summary of Changes During the Study Period

The bottom of Multimedia Appendix 2 includes three variables that summarize reported experiential changes over the course of the entire study: respondents reported an average of 1.83 new sexual partners, 53.9% (514/953) reported having sex without contraception at least once during the study, and 20.6% (196/953) reported a pregnancy during the study. We included these summary measures in our models presented in Multimedia Appendix 3 that used the 953 respondents as the units of analysis. Recall that we lost 39 of the original 992 respondents who only completed the first journal, and for whom we have no measures of change reported in the journal.

Time-Varying Pregnancy-Related Behaviors

We also explored the relationship of dynamic variables collected during the journal interviews to weekly survey participation. Table 1 provides means for the weekly interview questions about sex, partners, contraception, and pregnancy from the 57,602 journal interviews completed by the 953 respondents. Each question referred to the period since the prior journal (unless the interview was >14 days late, in which case it referred only to the prior week). As shown, most interviews indicated no changes from week to week, however, the following transitions were reported over the course of the study.

Table 1. Descriptive statistics for changes since prior journal (N=57,602 journals).

	Proportion/Mean
Sex	
No change	.84
Stopped having sex	.08
Started having sex	.08
Partner transitions	
No change	.90
Break-up (partner at time 1; no partner at time 2)	.04
New partner (no partner at time 1; partner at time 2)	.04
Partner switch (partner at time 1; different partner at time 2)	.03
Contraceptive Use	
No change	.92
Stopped using contraception	.04
Started using contraception	.04
Pregnancy	
No change	.99
Pregnancy ended	.005
New pregnancy	.004

An average of 7.65% (4408/57,602) of all journal interviews indicated that the respondent had stopped having sex since the prior interview (reported sex in the prior interview, but reported no sex in the current interview) and 7.62% (4389/57,602) reported that they started having sex (reported no sex in the prior interview, but reported sex in the current interview). In

terms of partner transitions, 4.08% (2353/57,602) of interviews indicated a break up (partner reported in the prior interview, but no partner reported in current interview), 4.08% (2353/57,602) indicated a new partner (no partner in the prior interview, partner in the current interview), and 2.69% (1548/57,602) indicated a partner switch (partner in the prior interview, different partner in the current interview). In terms of contraceptive use, 4.51% (2595/57,602) of interviews indicated stopping (contraceptive use in prior interview, no use in current interview) and 4.33% (2496/57,602) indicated starting (no contraceptive use in prior interview, use in current interview). Finally, 0.49% (284/57,602) of interviews indicated the end of a pregnancy (pregnant in prior interview, not pregnant in current interview) and 0.44% (254/57,602) indicated a new pregnancy (not pregnant in prior interview, pregnant in current interview).

We included these time-varying measures in our models that used the 57,602 journal interviews as the unit of observation (presented in Multimedia Appendix 4). Since each respondent will have experienced multiple changes between multiple pairs of journal interviews, we cannot include these measures of change in models using respondents as the unit of analysis (presented in Multimedia Appendix 3), but include instead the three summary measures of change from the bottom of Multimedia Appendix 2.

Dependent Variables and Analytic Methods

We used three dependent variables in our analyses: total days in the study for each respondent, total number of journal interviews completed by each respondent, and whether each journal interview was completed late.

Total days in study is a measure of attrition, and indicates the elapsed time from the first interview to the last interview. The mean days in study is 740.70, slightly more than two years, and ranges from 8 to 954 days. Figure 3 illustrates the distribution of this variable. Attrition is slightly higher in the first month, but declines steadily until 840 days (2.30 years). More than 63.4% (604/953) of the young women completed their final interview at 900 or more days (2.47 years). We used ordinary least squares (OLS) regression to model this dependent variable, with results presented in Multimedia Appendix 3.

The total number of journal interviews completed is a count of how many interviews the respondent completed. Among those who completed at least one journal interview after enrollment, the mean number of journals completed was 61.45, and ranges from 2 to 165. Approximately 15% of respondents completed fewer than 12 journal interviews. The target number of journals, which would require one every seven days, was 130. Figure 4 shows the distribution for this variable. As mentioned above, we encouraged respondents to complete journals every 7 days, but journals could be completed as early as 5 days after the previous journal was completed. Some respondents did choose to complete their journals more frequently than weekly and so they ended the study with more than 130 journals. We used OLS regression to estimate models of this dependent variable in Multimedia Appendix 3.

Finally, we looked at late completions, defining a *late* interview as one that occurred 14 or more days after the prior interview. An interview completed within 13 days still refers to changes since the prior interview, where at >14 days respondents were instructed to adjust the reference period to solely the week before, causing a period of missing data. As indicated in Table 2, only 11.07% (6376/57,602) of journals were completed late. We used logistic regression to estimate models of this dependent variable in Multimedia Appendix 4. This measure is specific to the journal interview, and it allows for an analysis of whether events that occurred just prior to the interview may have influenced the probability of late completion.

Table 2. Descriptive statistics for dependent variables

	Mean/ Proportion	SD	Range
Dependent Variables			
Total days in study (n=953)	740.70	303.68	8 – 954
Total number of journals completed (n=953)	61.45	42.54	2 – 165
Journal completed late (n=57,602)	.11	---	0 – 1

We used multilevel logistic regression (estimated using the SURVEYLOGISTIC procedure in Statistical Analysis System) to estimate the log-odds of the interview being late, which accounted for the multiple journal interviews for each respondent. This model takes the form

$$\text{logit}(p_{it}) = \beta_{0t} + \gamma * \mathbf{X}_i + \delta * \mathbf{Z}_{it}$$

where p_{it} is the probability of a journal interview being late, $\mathbf{X}_i$ is a set of non-time-varying characteristics of individual i, and $\mathbf{Z}_{it}$ is a set of time-varying characteristics associated with journal interview t, for each individual i. Thus, the model estimated effects on the probability of a particular journal interview being late as a function of both the characteristics of the individual completing the interview ($\mathbf{X}_i$) and the events that occurred since the prior journal interview ($\mathbf{Z}_{it}$).

Figure 3. Distribution of total days in study.
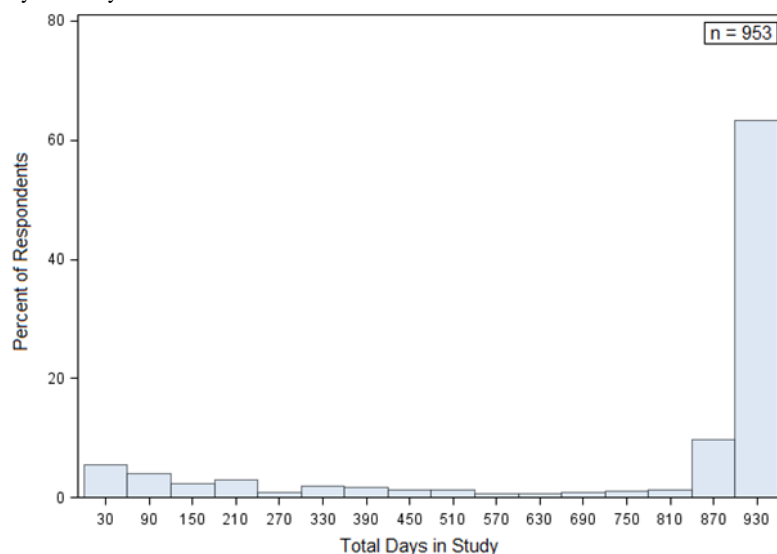
Figure 4. Distribution of total number of journals completed.
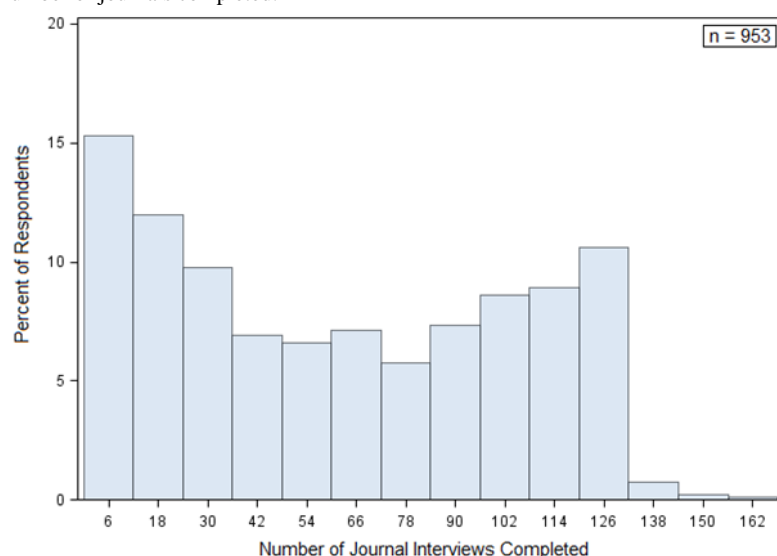
Results

[Multimedia Appendix 3](#) presents results for models predicting total days in the study (attrition) and total number of journal interviews completed. Results in columns 1 and 3 are for models using baseline measures only, while columns 2 and 4 are for models that add the three summary measures of experiences over the entire study period (number of new sexual partners, had sex without contraception, any pregnancy). These nested models allowed us to determine whether the relationship between baseline characteristics and participation were net of experiences during the study.

Total Days in Study

In the models of attrition shown in columns 1 and 2 of [Multimedia Appendix 3](#), positive numbers indicate more days in the study and negative numbers indicate fewer days in the study. Column 1 shows that African American women remained in the study for 45 fewer days than white women. Young women with less education, and/or a mother who was younger than 20

at her first birth, remained in the study for fewer days than their more advantaged peers. Highly religious respondents remained in the study longer than their less religious peers. Personality characteristics did not predict attrition. Respondents who provided both an email address and a phone number stayed in the study nearly four months (114 days) longer than those who provided only one or no type of contact information. Entering the study with a prior pregnancy was not associated with attrition, but respondents who had two or more prior pregnancies remained in the study for approximately 90 fewer days than those with no prior pregnancies.

The model shown in column 2 of [Multimedia Appendix 3](#), which includes the three summary measures of change, indicates that respondents who experienced changes during the study period remained in the study for a substantially longer period than their peers who experienced fewer or no changes. Respondents with more new sexual partners during the study period stayed in the study longer; on average, each new sex partner resulted in 27 additional days. Those who had sex without contraception stayed in the study for an average of 87 days longer than their peers

who did not have sex without contraception. Additionally, respondents who reported a pregnancy remained in the study for 141 days longer than those who did not report a pregnancy. The effects of sociodemographic background, personality characteristics, pregnancy-related experiences that occurred prior to the study, and contact information on the total number of days in the study are largely *net of* experiences during the study. The effects of experiences during the study are *net of* sociodemographics, personality, contact information, and prior experiences.

Note that because the dependent variable is not normally distributed, we also used logistic regression to model the probability of remaining in the study for at least 741 days (the mean value of the continuous variable). Results did not differ, so we present the OLS models due to their simpler interpretation.

Total Number of Completed Journals

Columns 3 and 4 in Multimedia Appendix 3 present OLS regression models of the total number of journal interviews completed by each respondent. This dependent variable represents both continued participation (total length of time in study is related to number of interviews completed) and timeliness (length of intervals between interviews is related to number of interviews completed). As such, it captures the quantity and quality of participation and, ultimately, the volume of information provided by respondents. In these models, positive coefficients indicate a larger number of completed journals at the end of the respondent's participation in the study, and negative coefficients indicate a smaller number of journals.

It is not surprising that some of the respondent characteristics that predict attrition were also associated with total number of journals completed, given the relationship between these two variables. In addition to predicting fewer total days in the study, the following characteristics also predicted completing fewer journals: being African American, completing high school but not currently enrolled in school, having a mother whose first birth was before age 20, not providing detailed contact information (both phone and email), and having risky adolescent experiences with sex and pregnancy (sex without contraception, and/or two or more pregnancies). However, column 3 shows that the following factors that predict total days in the study *were not* associated with total interviews completed: being enrolled in high school (relative to enrolled in a 4-year college), being highly religious, and growing up with a single parent. Conversely, having dropped out of high school (relative to enrolled in a 4-year college), growing up with an *other* parental living arrangement (relative to two parents), and extraversion all predicted completing fewer journal interviews, but had no association with total days in the study. Finally, conscientiousness was associated with completing almost 6 more interviews, but did not predict days in the study.

Column 4 of Multimedia Appendix 3 shows that changes in sex and pregnancy experienced during the study were strong predictors of the total number of completed journals. Respondents with more new sexual partners and pregnancy (or pregnancies) during the study period completed more journal interviews; on average, each new sex partner resulted in nearly 4 additional journals, and those who experienced a pregnancy completed almost 7 more journals than those who did not. As was the case for models of total days in the study, the effects of sociodemographics, personality, contact information, and prior experiences on the total number of journal interviews were largely *net of* experiences during the study, and the effects of experiences during the study were net of sociodemographics, personality, contact information, and prior experiences.

Late Journal Completion

Multimedia Appendix 4 presents models of whether each journal interview was completed late (>14 days since the prior interview). The unit of analysis is the journal interview rather than the respondent, therefore these models allowed us to examine if interviews were more likely to be late when changes in pregnancy-related behaviors were reported.

The coefficients in Multimedia Appendix 4 are the multiplicative effects on the odds of late journal completion (odds ratios). Thus, a coefficient of 1.00 indicates no effect, a coefficient less than 1.00 indicates lower odds of a late journal, and a coefficient greater than 1.00 indicates higher odds of a late journal. Column 1 presents results from the model based on respondent sociodemographic and personality characteristics, contact information/mode, and adolescent (pre-study) pregnancy-related experiences. The model shown in column 2 adds pregnancy-related changes since the prior journal.

A few background factors were associated with late journal completion: journals completed by African-American and/or less educated respondents were more frequently late than those completed by white/more-educated respondents. Journals completed by more conscientious respondents were less frequently late than those completed by less conscientious respondents. Respondents who provided both an email address and a phone number were less frequently late than those who provided only one or no type of contact information. Those who had sex without contraception during adolescence were more frequently late.

Pregnancy-related changes since the prior journal interview were strong and significant predictors of whether the subsequent journal interview was completed late. Weeks in which the respondent stopped having sex, started having sex, broke up with a partner, got a new partner, switched partners, stopped using contraception, started using contraception, got pregnant, or ended a pregnancy were all associated with late completion of the interview reporting these changes. Recall, however, that having these experiences throughout the study period was associated with less attrition from the study. Thus, although respondents with these experiences remained in the study longer, and completed more journal interviews, the interviews in which these behaviors were reported were more likely to be completed late.

Discussion

In sum, our findings on duration of participation in an intensive longitudinal study are consistent with more general findings on survey participation: minority and lower socioeconomic status respondents remained in the study for a shorter period than others [11,23,24], and also completed fewer journals and had

higher odds of late journal completion. Personality factors (extraversion, agreeableness, conscientiousness, neuroticism, and intellect/imagination) were not associated with overall time in the study, but extraverted and less conscientious respondents completed fewer interviews. Conscientious respondents also had lower odds of late journal completion. We found that contact information provided by the respondent was an important factor, with respondents who provided both an email address and a phone number remaining in the study longer, completed more journal interviews, and completed those interviews on time, compared to those who provided only one (or no) contact mode. This finding is consistent with existing research suggesting that pre-existing (at the time the study began) panel study commitment is a strong predictor of attrition [11]. We suspect that respondents who were more committed to participating in the RDSL study provided both email and phone contact information, while those less committed provided only one method, and those least committed provided no contact information.

Of greatest interest to our study, we found that experiences related to pregnancy were strong predictors of attrition and timeliness. Sex without contraception and two or more pregnancies before the study (during adolescence) were associated with higher levels of attrition and fewer total journal interviews during the study. Although these risky behaviors before the study were associated with attrition and fewer interviews, respondents with many changes in these behaviors during the study actually remained in the study longer and completed more interviews than their counterparts. This finding was reassuring since the RDSL was designed to investigate factors shaping the dynamics of sexual behavior, contraceptive use, and pregnancy. The specific weeks when changes in those risky sexual behaviors were reported, however, were more likely to be late (14 or more days since the prior interview) compared to weeks that continued the prior week's behavior. This trend may be consistent with other research finding that *shocks* (typically life events or negative experiences with the survey) can lead to attrition [33]. In this study, although these experiences did not seem to increase attrition, they did delay survey responses. These results may also be consistent with research finding that health status predicts attrition [24], although that research did not specifically address pregnancy as a health status.

Our finding that respondents with dynamic experiences that were measured in the surveys remained in the study longer and completed more journal interviews is consistent with studies finding that boring survey modules increase attrition [11], and that tailored questionnaire content reduces attrition [21]. Further, we know from semi-structured interviews and an open-ended question at the end of each weekly journal interview that some respondents found the long-term and repetitive nature of the journal interviews tedious. This problem is highlighted by several responses to the weekly question, "Is there anything else you would like to tell us?" Responses included, "I think that you guys need to ask different questions. I am answered (sic) the same questions over and over again.", "No but I think you should make the survey shorter or combine it to a couple of pages.", and "Same questions every week!!! Come on."

We speculate, therefore, that perhaps the frequent interviews may have seemed less burdensome for respondents whose lives included more changes in the experiences that were the focus of the study. Changes in these experiences would result in more varied survey content from week to week. This issue was also hinted at in some of the responses to the open-ended question at the conclusion of the journal, for example, "Doing this survey helps me to vent quite a bit because I don't have many people around me that I can tell the whole truth to. So thank you for this opportunity…"

Overall, our analyses suggest that attrition was a relatively smooth process in our study, with slightly higher rates in the first month, but then low rates that increased slowly and steadily until the end of the study. Further, the vast majority of respondents (90.7%, 865/953) completed at least half of their journal interviews on time (before 14 days had passed).

Our study focused on a young age group (18 and 19 years old at the beginning of the study, and 20 to 22 years at the end) that typically has higher attrition than older adults [23], suggesting that our attrition rate may have been higher than similar studies of older age groups would be. However, given that tailored content can reduce attrition, and our substantive topics were so important to this age group, our study may have had lower attrition than other studies. We present this as a case study, with the hope that further studies will explore attrition in surveys about other topics, and with other age groups.

Limitations and Future Research

This study of continued and on-time participation in a weekly survey, as well as the RDSL study more generally, has several important limitations. Most significant, we lack crucial information about respondents who dropped out of the study. Although we can summarize changes experienced during the period that attriters remained in the study, we do not know what events occurred just prior to their decisions to stop participating. For example, although we found that respondents dropped out at higher rates after reporting a pregnancy, some respondents may have left the study when they discovered they were pregnant but had not yet reported it. We do not believe that this is the case, considering our study's pregnancy rates closely resemble the vital statistics rates for this age group in this area, but we cannot rule out this possibility. Further, we do not know what other unmeasured experiences may have contributed to respondents' decisions to drop out.

The narrow geographic focus (a single county in Michigan) of the RDSL study is also a limitation. However, although the sample is not representative of the United States as a whole, Michigan falls close to the national median for many of the measures of interest in this study: cohabitation, marriage, age at first birth, completed family size, non-marital childbearing, and teenage childbearing [34]. This information does not suggest that Michigan is representative of the nation, but implies that it is not an outlier. Another constraint is that the study includes only a small number of Latinas, which is a growing demographic group in the United States. We hope that the research findings of the RDSL will motivate future researchers to implement journal methods on larger and more diverse populations.

This study focused on a narrow age range, and a set of topics that were of great importance to the study population. It is difficult to know how these general conclusions about attrition – that those with the most experiences related to the subject of the study participate longer and complete more interviews, but have higher odds of late interviews – would translate to surveys on other age groups or on other topics. It is possible that studies focused on topics of specific interest to the age group in the study might have less attrition than others.

Finally, our study design did not include experiments related to reducing attrition. With a few minor differences, all respondents were subject to the reminder protocol described in Figure 2, and the incentive protocol described above, across all weeks in the study. Additionally, all respondents were asked to complete the surveys weekly and so we do not know if participation would have been different if we had asked some or all respondents to complete the surveys less frequently. Future intensive longitudinal surveys should include experiments designed to test and optimize reminder and incentive protocols, and to determine the ideal frequency of measurement.

Experiments are also needed to determine how to keep respondents engaged in intensive longitudinal surveys. In this study, the complicated nature of the association between pregnancy-related behaviors and study participation – the reduced timeliness of interviews during weeks of change, but the larger number of interviews for respondents who experienced more changes in these behaviors – highlights the need for responsive designs in social research. For example, respondents who are not sexually active might be more engaged if the frequency of interviews was reduced to monthly, or if interview content could be tailored to better reflect their experiences. Respondents who are sexually active, and who may have trouble completing their journal interviews during particularly busy weeks of their lives, may find it easier to participate via a mobile version of the journal interview. In fact, if we were to conduct this study again today, we would have a mobile site or app available to complete the survey, given current knowledge and the increase in smartphone and mobile device use among this population [35].

Conclusions

We have demonstrated in this paper that participation in an intensive longitudinal study with weekly web surveys may vary not only by respondent characteristics, but also by behaviors measured across the surveys. Our analyses suggest that respondents who experience the behaviors measured by the study may participate more than respondents who do not experience those behaviors. These results highlight the need for experiments to determine the most effective data collection strategies and procedures for assuring continued and on-time participation in intensive longitudinal studies.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Sample journal interviews.

[PDF File (Adobe PDF File), 355KB - jmir_v18i6e105_app1.pdf]

Multimedia Appendix 2

Descriptive statistics for individual characteristics (N=953).

[PDF File (Adobe PDF File), 18KB - jmir_v18i6e105_app2.pdf]

Multimedia Appendix 3

OLS regression models of total days in study and total number of journals completed (n=953 respondents).

[PDF File (Adobe PDF File), 38KB - jmir_v18i6e105_app3.pdf]

Multimedia Appendix 4

Logistic regression models of late journal completion (coefficients are odds ratios; confidence intervals in parentheses; n=953 respondents; n=57,602 journal interviews).

[PDF File (Adobe PDF File), 32KB - jmir_v18i6e105_app4.pdf]

References

1. Walls T, Schafer J. Models for intensive longitudinal data. Oxford, UK: Oxford University Press; 2006.
2. Larson R, Csikszentmihalyi M. The experience sampling method. New Dir Methodol Soc Behav Sci. 1983. p. 41-56 URL: http://www.springer.com/cda/content/document/cda_downloaddocument/9789401790871-c2.pdf?SGWID=0-0-45-1481923-p176723966 [WebCite Cache ID 6hBvWJKk0]
3. Denscombe M. Item nonresponse rates: a comparison of online and paper questionaires. Int J Soc Res Methodol. 2009 Oct. (4) p. 281-291 URL: https://www.researchgate.net/publication/248988677_Item_non-response_rates_A_comparison_of_online_and_paper_questionnaires [WebCite Cache ID 6iJOc9cut]
4. Fox J, Murray C, Warm A. Conducting research using web-based questionnaires: practical, methodological, and ethical considerations. Int J Soc Res Methodol. 2003. (2) p. 167-180 URL: http://sociology.soc.uoc.gr/socmedia/papageo/web-based%20questionnaires.pdf [WebCite Cache ID 6iJOgk4ty]
5. Couper M, Miller P. Web survey methods introduction. Public Opin Q. 2008 Dec 1. (5) p. 831-835 URL: http://poq.oxfordjournals.org/content/72/5/831.full.pdf+html [WebCite Cache ID 6iJOjeZm4]
6. Michie S, Marteau T. Non-response bias in prospective studies of patients and health care professionals. Int J Soc Res Methodol. 1999. (3) p. 203-212 URL: http://www.tandfonline.com/doi/pdf/10.1080/136455799295014 [WebCite Cache ID 6iJOp5lS7]
7. Behr A, Bellgardt E, Rendtel U. Extent and determinants of panel attrition in the European Community Household Panel. Eur Sociol Rev. 2005 Dec 1. (5) p. 489-512 URL: http://esr.oxfordjournals.org/content/21/5/489.full.pdf+html [WebCite Cache ID 6iJOuMyYC]
8. Fitzgerald J, Gottschalk P, Moffitt R. An analysis of sample attrition in panel data: the Michigan panel study of income dynamics. J Hum Resour 1998. 1998 Apr 1. (2) p. 251-299 URL: http://personal.psc.isr.umich.edu/yuxie-web/files/pubs/Articles/Fitzgerald_et_al1998.pdf [WebCite Cache ID 6iJOrlAB7]
9. Lillard L, Panis C. Panel attrition from the Panel Study of Income Dynamics: household income, marital status, and mortality. J Hum Resour. 1998 Apr 1. (2) p. 437-457 URL: http://www.jstor.org/stable/146436?seq=1#page_scan_tab_contents [WebCite Cache ID 6iJOxtQ6e]
10. Loosveldt G, Pickery J, Billiet J. Item nonresponse as a predictor of unit nonresponse in a panel survey. J Off Stat-Stockh. 2002. (4) p. 545-558 URL: http://www.jos.nu/Articles/article.asp [WebCite Cache ID 6iJPFCnsw]
11. Lugtig P. Panel attrition: separating stayers, fast attriters, gradual attriters, and lurkers. Sociol Methods Res. 2014 Nov 1. (4) p. 699-723 URL: http://smr.sagepub.com/content/early/2014/02/06/0049124113520305.full.pdf+html [WebCite Cache ID 6iJPJt8Eh]
12. Nicoletti C, Peracchi F. Survey response and survey characteristics: microlevel evidence from the European Community Household Panel. J R Stat Soc Ser A Stat Soc 2005 Jan 1;168(4):763-781.
13. Reinwand DA, Crutzen R, Elfeddali I, Schneider F, Schulz DN, Stanczyk NE, et al. Impact of educational level on study attrition and evaluation of web-based computer-tailored interventions: results from seven randomized controlled trials. J Med Internet Res 2015 Oct 7;17(10):e228 [FREE Full text] [doi: 10.2196/jmir.4941] [Medline: 26446779]
14. Trappmann M, Gramlich T, Mosthaf A. The effect of events between waves on panel attrition. Surv Res Methods 2015 Apr 10;9(1):31-43.
15. American Association of Public Opinion Research. Standard definitions: final dispositions of case codes and outcome rates for surveys. Oakbrook Terrace: AAPOR; 2011. URL: https://www.esomar.org/uploads/public/knowledge-and-standards/codes-and-guidelines/ESOMAR_Standard-Definitions-Final-Dispositions-of-Case-Codes-and-Outcome-Rates-for-Surveys.pdf [accessed 2016-04-20] [WebCite Cache ID 6hBvWJKk0]
16. Gatny H, Couper M, Axinn W, Barber J. Using debit cards for incentive payments: experiences of a weekly survey study. Surv Pract 2009;2(7):1-5.
17. Crawford S, Couper M, Lamias M. Web Surveys: Perceptions of Burden. Social Science Computer Review 2001 May 01;19(2):146-162. [doi: 10.1177/089443930101900202]
18. Kaplowitz M, Hadlock T, Levine R. A Comparison of Web and Mail Survey Response Rates. Public Opinion Quarterly 2004 Apr 22;68(1):94-101. [doi: 10.1093/poq/nfh006]
19. Porter SR, Whitcomb ME. Mixed-Mode Contacts In Web Surveys: Paper is Not Necessarily Better. Public Opinion Quarterly 2007 Sep 18;71(4):635-648. [doi: 10.1093/poq/nfm038]
20. Kittleson MJ. Determining effective follow-up of email surveys. Am J Health Behav 1997;21(3):193-196.
21. Oudejans M. Can tailor-made questions motivate respondents in an internet panel? 2011 Presented at: European survey research associaton conference; 2011; Lausanne, Switzerland.

22. Barber JS, Kusunoki Y, Gatny HH. Design and implementation of an online weekly journal to study unintended pregnancies. Vienna Yearb Popul Res 2011 Jan 1;9:327-334 [FREE Full text] [Medline: 22408644]

23. Lipps O. Attrition of households and individuals in panel surveys. In: Lynn P, editor. Methodology of longitudinal surveys. Chichester, UK: John Wiley & Sons; 2009.

24. Watson N, Wooden M. Identifying factors affecting longitudinal survey response. In: Lynn P, editor. Methodology of longitudinal surveys. Chichester, UK: John Wiley & Sons; 2009.

25. Axinn WG, Gatny HH, Wagner J. Maximizing data quality using mode switching in mixed-device survey design: nonresponse bias and models of demographic behavior. Methoden Daten Anal 2015;9(2):164-184. [Medline: 26865882]

26. Barber JS, Gatny HH, Kusunoki Y, Schulz P. Effects of intensive longitudinal data collection on pregnancy and contraceptive use. Int J Soc Res Methodol 2014 Dec 1;19(2):205-222.

27. Barber JS, Yarger JE, Gatny HH. Black-white differences in attitudes related to pregnancy among young women. Demography 2015 Jun;52(3):751-786 [FREE Full text] [doi: 10.1007/s13524-015-0391-4] [Medline: 25962867]

28. Gatny HH, Kusunoki Y, Barber JS. Pregnancy scares and subsequent unintended pregnancy. Demogr Res 2014 Nov 20;31:1229-1242 [FREE Full text] [doi: 10.4054/DemRes.2014.31.40] [Medline: 25642141]

29. Hall KS, Kusunoki Y, Gatny H, Barber J. The risk of unintended pregnancy among young women with mental health symptoms. Soc Sci Med 2014 Jan;100:62-71 [FREE Full text] [doi: 10.1016/j.socscimed.2013.10.037] [Medline: 24444840]

30. Hall KS, Moreau C, Trussell J, Barber J. Role of young women's depression and stress symptoms in their weekly use and nonuse of contraceptive methods. J Adolesc Health 2013 Aug;53(2):241-248 [FREE Full text] [doi: 10.1016/j.jadohealth.2013.02.009] [Medline: 23582524]

31. Miller WB, Barber JS, Gatny HH. The effects of ambivalent fertility desires on pregnancy risk in young women in the USA. Popul Stud (Camb) 2013;67(1):25-38 [FREE Full text] [doi: 10.1080/00324728.2012.738823] [Medline: 23234316]

32. Costa P, McCrae R. Multiple uses for longitudinal personality data. Eur J Personal 1992 Jun 1;6(2):85-102.

33. Uhrig S. ISER Working Paper Series. 2008. The nature and causes of attrition in the British Household Panel Study URL: http://www.econstor.eu/handle/10419/92025 [accessed 2016-04-20] [WebCite Cache ID 6hBw1kxd8]

34. Lesthaeghe R, Neidert L. The second demographic transition in the United States: exception or textbook example? Popul Dev Rev 2006 Dec 1;32(4):669-698.

35. Moscaritolo A. Surprise: many teens are online 'almost constantly'. 2015 Apr 09. URL: http://www.pcmag.com/article2/0,2817,2480788,00.asp [accessed 2016-04-04] [WebCite Cache ID 6hBwCMpgr]

Abbreviations

NEO-FFI: Neuroticism, Extraversion, and Openness to Experience-Five Factor Inventory
OLS: Ordinary least squares
OR: odds ratio
PID: Personal Identification Card
RDSL: Relationship Dynamics and Social Life

Original Paper

Achieving Synergy: Linking an Internet-Based Inflammatory Bowel Disease Cohort to a Community-Based Inception Cohort and Multicentered Cohort in Inflammatory Bowel Disease

Bharati Kochar[1], M.D.; Molly Aldridge[2], MPH; Suzanne Follan Cook[2], PhD; Renee Bright[3], MS; Meaghan Mallette[3], MPA; Heather Moniz[3], BS; Samir A Shah[4], MD; Neal S LeLeiko[4,5], MD, PhD; Jason Shapiro[4,5], MD; Bruce E Sands[6], MD, MS; Wenli Chen[1], MA, MS; Elizabeth Jaeger[1]; Joseph Galanko[1], PhD; Millie D Long[1], MD, MPH; Christopher F Martin[1], MSPH; Robert S Sandler[1], MD, MPH; Michael D Kappelman[1], MD, MPH

[1]University of North Carolina, Center for Gastrointestinal Biology and Disease, Chapel Hill, NC, United States
[2]GlaxoSmithKline, Research Triangle Park, NC, United States
[3]Rhode Island Hospital, Providence, RI, United States
[4]The Warren Alpert Medical School at Brown University, Providence, RI, United States
[5]Hasbro Childrens' Hospital, Division of Pediatric Gastroenterology, Nutrition and Liver Diseases, Providence, RI, United States
[6]Mt. Sinai Medical Center, Division of Gastroenterology, New York City, NY, United States

Corresponding Author:
Bharati Kochar, M.D.
University of North Carolina
Center for Gastrointestinal Biology and Disease
Campus Box #7080, Bioinformatics Building
130 Mason Farm Road
Chapel Hill, NC, 27599
United States
Phone: 1 9199662514
Fax: 1 9199666842
Email: bharati.kochar@unchealth.unc.edu

Abstract

Background: Traditional cohort studies are important contributors to our understanding of inflammatory bowel diseases, but they are labor intensive and often do not focus on patient-reported outcomes. Internet-based studies provide new opportunities to study patient-reported outcomes and can be efficiently implemented and scaled. If a traditional cohort study was linked to an Internet-based study, both studies could benefit from added synergy. Existing cohort studies provide an opportunity to develop and test processes for cohort linkage. The Crohn's and Colitis Foundation of America's (CCFA) Partners study is an Internet-based cohort of more than 14,000 participants. The Ocean State Crohn's and Colitis Area Registry (OSCCAR) is an inception cohort. The Sinai-Helmsley Alliance for Research Excellence (SHARE) is a multicentered cohort of inflammatory bowel disease patients. Both the later cohorts include medical record abstraction, patient surveys, and biospecimen collection.

Objective: Given the complementary nature of these existing cohorts, we sought to corecruit and link data.

Methods: Eligible OSCCAR and SHARE participants were invited to join the CCFA Partners study and provide consent for data sharing between the 2 cohorts. After informed consent, participants were directed to the CCFA Partners website to complete enrollment and a baseline Web-based survey. Participants were linked across the 2 cohorts by the matching of an email address. We compared demographic and clinical characteristics between OSCCAR and SHARE participants who did and did not enroll in CCFA Partners and the data linkage.

Results: Of 408 participants in the OSCCAR cohort, 320 were eligible for participation in the CCFA Partners cohort. Of these participants, 243 consented to participation; however, only 44 enrolled in CCFA Partners and completed the linkage. OSCCAR participants who enrolled in CCFA Partners were better educated (17% with doctoral degrees) than those who did not (3% with doctoral degrees, *P*=.01). In the SHARE cohort, 436 participants enrolled and linked to the Partners cohort. More women (60% vs 50%) linked and those who linked were predominantly white (96%; *P*<.01). Crohn's disease patients who linked had lower

XSL•FO
RenderX

mean scores on the Harvey-Bradshaw Index (3.6 vs 4.4, *P*<.01). Ulcerative colitis patients who linked had less extensive disease than those who did not link (45% vs 60%, *P*<.01).

Conclusions: Linkage of CCFA Partners with cohorts such as OSCCAR and SHARE may be a cost-effective way to expand the infrastructure for clinical outcomes and translational research. Although linkage is feasible from a technical, legal, and regulatory perspective, participant willingness appears to be a limiting factor. Overcoming this barrier will be needed to generate meaningful sample sizes to conduct studies of biomarkers, natural history, and clinical effectiveness using linked data.

Introduction

Cohort studies are designed to evaluate risk factors for developing a health condition and/or factors which affect the natural history of disease [1]. In traditional cohort studies, subjects are extensively characterized at baseline (eg health history, environmental exposures), often with the collection of biosamples and observed over time for the development of outcomes of interest. These outcomes usually include important medical events such as surgery, cancer, or death. In recent years, there has been growing interest in the use of the Internet and social media to develop electronic cohorts, which focus on patient-reported exposures and health behaviors and patient-reported outcomes (PROs) such as fatigue, well-being, and illness experience [2-5]. Electronic cohorts have the advantage of efficient recruitment and data collection and may be less costly to maintain over time. Linking a traditional cohort with an electronic cohort offers an opportunity to derive the benefits of both cohort designs. This hybrid cohort would have the benefit of well-characterized, validated cases and biological specimens. The speed and low cost of the electronic component would provide a cost-effective approach to evaluate an extended array of PROs. Translational studies could exploit the biological samples in the hybrid cohort.

Two novel and complementary cohort studies initiated by the Crohn's and Colitis Foundation of America (CCFA) provide an opportunity to develop and test the process of cohort linkage. The Ocean State Crohn's and Colitis Area Registry (OSCCAR), a traditional cohort study, is an inception cohort that includes medical record abstraction, patient surveys, and biospecimen collection. CCFA Partners is an electronic cohort of more than 14,000 participants with inflammatory bowel diseases (IBDs) that focuses on patient-reported exposures and outcomes. In addition, IBD patients enrolled in the Sinai-Helmsley Alliance for Research Excellence (SHARE), a 7-center prospective observation cohort, were also evaluated to test the linkage of a traditional cohort to an electronic cohort. As these cohorts have different areas of focus, they are complementary to one another. If successful, linkage would provide a cost-effective strategy to expand the data assets of both cohorts, enabling studies not possible in either cohort alone. Record linkages of this kind have the potential to create information-rich environments [6] allowing for a more patient-centered and holistic approach to observational research, patient care, and improved patient outcomes.

Other record linkage initiatives have been established around the world for a variety of research purposes. For example, Australia's Population Health Research Network is expected to be the world's "largest population database supporting health research, policy, and planning" and links together a number of health administrative databases [7]. Researchers in Canada are working to link health administration data with data from the Canadian Longitudinal Study on Aging cohort and the Canadian Partnership for Tomorrow Project cohort [8]. An initiative in the United Kingdom aimed to link PRO measures reported through an Web-based data collection system to clinical data in a cancer registry [9]. To our knowledge, our study represents the first published attempt to link data from a traditional, inception cohort to a large electronic cohort.

The objectives of this project were (1) to develop the technical, regulatory, and legal infrastructure to corecruit OSCCAR and SHARE participants and into CCFA Partners and link the data assets of these 2 cohorts and (2) to evaluate the feasibility of this corecruitment and linkage process. If successful, linkage of these cohorts would produce a "virtual database" of expanded clinical, PROs, and biospecimen data.

Methods

Study Setting

OSCCAR

OSCCAR is a novel, community-based inception cohort of 408 Crohn's disease (CD) and ulcerative colitis (UC; IBD subtypes) patients launched in Rhode Island in 2008 and closed to enrollment in 2013. Funded by the Centers for Disease Control and Prevention, investigators at Mount Sinai Hospital, Harvard Medical School, and the Warren Alpert Medical School of Brown University jointly manage the OSCCAR cohort. The cohort is designed to collect data on incidence rates of CD and UC, disease outcomes, and factors predictive of disease onset and outcomes. In addition to the prospective and longitudinal collection of clinical data by physician report and chart abstraction, OSCCAR participants complete regular surveys and contribute to a biobank of blood, urine, and stool specimens [10].

Newly diagnosed pediatric and adult patients with IBD who were residents of Rhode Island were referred by their diagnosing gastroenterologist or surgeon for enrollment in OSCCAR from January 2008 to January 2013. After informed consent, diagnosis of IBD was confirmed by chart review by medically trained

study personnel according to standard criteria of the National Institutes of Diabetes and Digestive and Kidney Diseases IBD Genetic Consortium. Participants were visited at study enrollment and requested to provide blood, urine, and stool samples. All subjects were also asked to complete the Inflammatory Bowel Disease Questionnaire, IMPACT Questionnaire, Short Form Health Survey (SF)-36, EuroQoL 5, Work Productivity and Activity Impairment Questionnaire, Functional Assessment of Chronic Illness Therapy, Short Inflammatory Bowel Disease Questionnaire, An Assessment of Environmental Risks for Ulcerative Colitis and Crohn's Disease in the United States, and PHQ-8 Depression Diagnostic and Severity Measure [11-17]. Subjects with CD were asked to complete the Harvey-Bradshaw Index (HBI). Subjects with UC were asked to complete the Simple Clinical Colitis Activity Index (SCCAI). The participants are contacted every 6 months through the study period to update their profiles and collect specimens [10, 18].

Share

The SHARE is a prospective observational cohort of patients with IBDs. The cohort was established with the goal of creating a database of clinical information and a repository of biological specimens for genetic, molecular, and microbiological research to better understand IBD and to help develop new therapies. Starting in July 2012, patients were recruited from 7 academic centers around the United States: Mount Sinai Hospital, University of Chicago Medical Center, Massachusetts General Hospital, Cedars-Sinai Medical Center, Mayo Clinic, Rochester, University of North Carolina at Chapel Hill Hospital, and Washington University, Saint Louis, Medical Center. Enrollment is currently ongoing.

Patients who are aged younger than 18 years, unable to understand or provide informed consent, and do not have a confirmed diagnosis of IBD in their medical records were excluded. Recruited patients provided demographic information, medical history, surgical history, family history, medication use, extraintestinal manifestations, and blood samples via an interview with a study coordinator. Disease characteristics, medication use, extraintestinal manifestations, and laboratory results are obtained from the medical record by the study coordinator. The HBI, SCCAI, and Modified Pouchitis Disease Activity Index are completed during the study visit by the investigator or a study coordinator [19-21]. The Depression Activity and Severity Index (PHQ-8) and Manitoba Index are completed by the patient during the study visit [17, 22].

CCFA Partners

CCFA Partners is a validated, Internet-based, long-term cohort study of adult subjects with IBD initiated in July 2011 and administered by the University of North Carolina, Chapel Hill. The development of CCFA Partners has been described in detail previously [23]. In brief, participants with a self-reported diagnosis of UC, CD, or indeterminate colitis were recruited through CCFA email rosters, the CCFA website, social media outlets, and at educational and fundraising events. The baseline and subsequent surveys administered every 6 months include questions regarding demographics, disease type, disease course and activity, IBD-related treatments, and concurrent medication

use including alternative medicines, family history of disease, diet, exercise, and PROs including quality of life and medication adherence [23]. To date, over 14,000 people with IBD have been enrolled. Published and planned studies using data generated from the CCFA Partners e-cohort include the evaluation of (1) the impact of IBD on important, patient-centered outcomes, (2) the role of environmental influences (eg, sleep, depression, diet, exercise) on disease activity, (3) patient perceptions on selected topics such as quality of care, biobanking, and so forth, and (4) clinical effectiveness and drug safety research [24-29]. Further details on the validation of patient-reported data in a subset of CCFA Partners are published [30].

Linkage Process

Eligibility

To enable record linkage, we identified members of OSCCAR and SHARE who met the criteria for membership in CCFA Partners, including having a diagnosis of IBD and access to the Internet.

Education and Recruitment

OSCCAR participants are seen in person at years 1 and 5 after enrollment and are contacted by phone every 6 months. During these points of contact, OSCCAR staff provided participants with recruitment materials about the CCFA Partners cohort and the OSCCAR-CCFA Partners linkage study. Participants were also provided with the CCFA Partners Web address for additional information. All OSCCAR participants were contacted on at least three occasions.

SHARE participants were asked if they were interested in participating in CCFA Partners at the time of their enrollment in SHARE between July 2012 and December 2014. Cedars-Sinai Medical Center did not participate in this linkage study.

Linkage

Participants were linked across the 2 cohorts by the matching of their email addresses. After informed consent for the data linkage, participants were sent an automated email with instructions on how to access the CCFA Partners website to complete CCFA Partners enrollment/consent and to fill out the baseline survey. Each participant was sent up to three email reminders. In addition, in the OSCCAR cohort, the study coordinator and research assistants sent personalized letters and made phone calls to each individual to encourage completion of the linkage process.

Ethics

The study protocol was approved by the Institutional Review Boards at Rhode Island Hospital, the University of North Carolina at Chapel Hill, and the Institutional Review Boards of the SHARE sites that participated. Personal health information was protected across linkages by using deidentified study identification numbers. Where necessary, personal identifiers for consented participants were made available to trained research staff identifiers through a password-protected Web portal.

XSL•FO

RenderX

Analysis

The enrolled subjects were compared to the unenrolled subjects. Categorical variables were compared using Fisher's exact tests or chi-square tests as appropriate and continuous variables were compared using t tests. All P values less than .05 were considered statistically significant. Analyses were performed using SAS 9.4 (SAS Institute, Cary, NC) and Stata 14.0 (StataCorp, College Station, TX).

Results

OSCCAR

Figure 1 illustrates the recruitment process for the CCFA Partners/OSCCAR linkage. A total of 408 participants have been enrolled in OSCCAR. Seventy-six of these subjects are no longer being followed because of voluntary withdrawal, loss to follow-up, criteria insufficient for a diagnosis of IBD, or death. An additional 12 subjects report not having access to the Internet. Of the remaining 320 participants who were eligible for linkage, all were approached for participation. Two hundred forty-three subjects consented to linkage with the Partners cohort, 25 declined participation, and the remaining 52 did not

respond to recruitment efforts. Of those who consented to participation, 44 (14% of those eligible) completed the linkage.

Table 1 shows characteristics of OSCCAR cohort members who were successfully enrolled in and completed linkage with CCFA Partners compared with members who did not. The OSCCAR cohort is predominantly white (97%) and race did not appear to influence CCFA Partners enrollment and linkage. The mean age at diagnosis was 34 years in the linked group and 33 years in the unlinked groups. Similarly, disease type did not affect enrollment. Of those who completed linkage, 28 (63%) had CD and of those who did not 192 (60%) had CD. Fourteen (32%) and 122 (38%) of linked and unlinked subjects were diagnosed with UC, respectively.

OSCCAR participants who were successfully recruited into CCFA Partners were more educated (8 participants [17%] with a doctoral degree) than those who were not (7 participants [3%] with a doctoral degree, $P=.01$). Similarly, those who did not enroll in CCFA Partners were numerically more likely to have not attained a bachelor's degree than those who did enroll; however, this was not statistically significant. There were no other significant demographic or clinical differences (eg, disease activity, disease behavior, disease location, medication use).

Figure 1. Recruitment efforts for the CCFA Partners/OSCCAR record linkage.
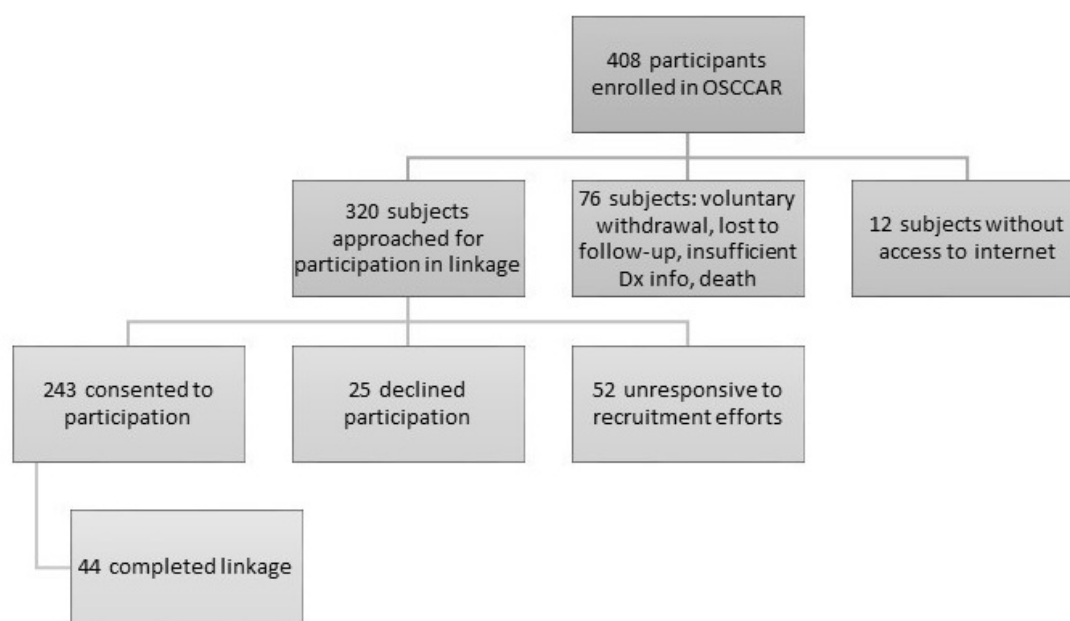
Table 1. Characteristics of OSCCAR cohort members who linked with CCFA Partners

			Linked (n=44)	Unlinked (n=224)	*P* value
Demographics					
	% Female		75	64.3	.22
	Race				1
		% White	97.7	96.6	
		% Black	2.3	3.4	
		% Hispanic or Latino	4.6	5.8	
		% Ashkenazi Jewish	11.9	4.2	.06
	Education				
		% Less than Bachelor's degree	34.5	53.2	.07
		% Bachelor's degree	41.4	31.4	.29
		% Master's degree	6.9	12.2	.54
		% Doctoral degree	17.2	3.2	.01
	Occupation				.78
		% Paid occupation	54.6	50	
		% Disabled	2.3	4.9	
		% Other	43.2	45.1	
	Married		25	36.3	.17
	Ever smoked		18.2	17	.83
Disease course					
	Mean age at diagnosis (years)		34	33	.89
	Diagnosis				.39
		% Diagnosis of Crohn's disease	63.4	59.8	
		% Diagnosis of ulcerative colitis	31.7	37.9	
		% Diagnosis of indeterminate colitis	4.9	2.3	
	% History of bowel resections		6.5	16.7	.18
	% Pyoderma gangrenosum		0	0	1
	% Erythema nodosum		0	0	1
	% Scleritis/episcleritis/iritis/uveitis		4.6	4	1
	% Sacroiliitis		4.6	1.3	.19
	% Family history of IBD		32	35	.73
Medications					
	% Steroid use[a]		4.9	3.7	.66
	% Aminosalicylate use[a]		41.5	37.5	.73
	% Immunomodulators[a]		12.2	19.4	.38
	% Biologic use[a]		24.4	25	1
Crohn's disease					1
	% Ileal		20	21.5	
	% Colonic		28	27.7	
	% Ileo-colonic		52	50.1	

		Linked (n=44)	Unlinked (n=224)	*P* value
	% Concomitant upper disease	15.4	20.6	.79
% Fistulae		2.3	0.9	.42
Behavior				1
	% Inflammatory	88	83.2	
	% Stricturing	8	9.9	
	% Penetrating	4	6.9	
% Perianal disease		11.5	7.6	.85
Mean Harvey-Bradshaw Index[a]		2.3	2.5	.56
Ulcerative colitis				1
	% Proctitis	15.4	19.3	
	% Left-sided colitis	23.1	26.5	
	% Extensive colitis	61.5	54.2	
Mean Clinical Colitis Activity Index[a]		3.3	2.3	.17

[a]At the time of contact

Share

Of the SHARE participants from July 2012 to December 2014 who were approached to enroll in the CCFA Partners cohort, 1671 (75%) consented to linkage. However, only 24% (n=436) completed the linkage by enrolling in the Partners cohort (see Figure 2). Enrollment varied by study center. Table 2 shows characteristics of SHARE cohort members who were successfully enrolled in and completed linkage with CCFA Partners compared with members who did not. Gender was a notable difference: of those who linked, 263 (60%) were women, compared to 898 (50%) of those who did not link (*P*<.01). Linked subjects were predominantly white (96%; n=417); there were more blacks who did not link (6%; n=115; *P*<.01). Of those who linked, 49% (215) were using biological agents compared to 43% (767) of those who did not enroll (*P*=.02). There were no significant differences in patterns of use for other medication classes. Although there were no significant differences in disease location in CD, linked subjects had a mean HBI of 3.6 and those who did not link had a mean HBI of 4.4 (*P*<.01). Among those subjects with UC, 40% (176) of linked subjects had left-sided colitis compared to 27% (491) of those who did not link and 45% (196) of linked subjects had extensive colitis, compared to 60% (1066) of those who did not link (*P*<.01). However, there was no difference in the mean SCCAI score.

Figure 2. Recruitment efforts for the CCFA Partners/SHARE record linkage.
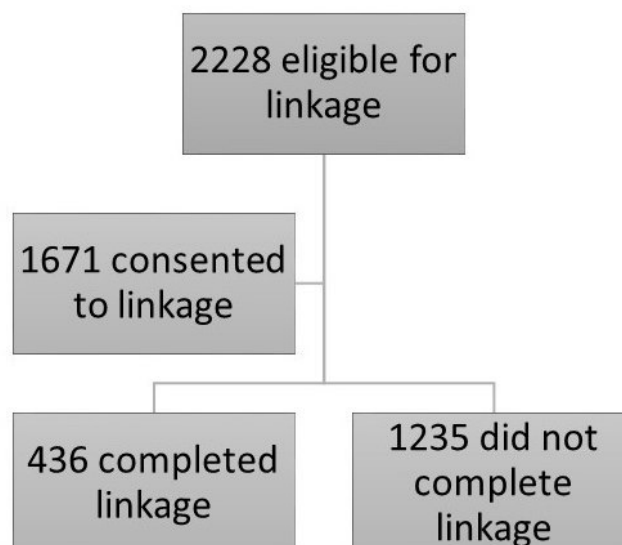
Table 2. Characteristics of SHARE cohort members who linked with CCFA Partners.

			Linked (n=436)	Unlinked (n=1792)	*P* value
Demographics					
	Mean age (years)		39	40	.10
	% Female		60.4	50.1	<.01
	Race				<.01
		% White	95.6	86.8	
		% Black	1.2	6.4	
		% Latino	2.8	3.5	.48
		% Ashkenazi Jewish	17.8	12.2	.15
	Ever smoked		29.3	34.2	.07
Disease course					
	Mean age at diagnosis (years)		29	29	.29
	Diagnosis				.20
		% Diagnosis of Crohn's disease	65.4	60.8	
		% Diagnosis of ulcerative colitis	32.6	36.7	
		% Diagnosis of indeterminate colitis	2.1	2.6	
	% History of bowel resections		32.1	33.9	.48
	% Pyoderma gangrenosum		0.9	2.7	.03
	% Erythema nodosum		5.6	3.4	.03
	% Scleritis/episcleritis/iritis/uveitis		5.8	3.6	.04
	% Sacroiliitis		2.1	1.4	.28
	% Family history of IBD		24.5	22.6	.43
Medications					
	% Steroid use[a]		10.0	10.4	.83
	% Aminosalicylate use[a]		32.3	35.4	.23
	% Immunomodulators[a]		39.1	38.1	.70
	% Biological use[a]		49.3	42.8	.02
Crohn's disease					.47
	% Ileal		29.3	25.7	
	% Colonic		20.4	20.8	
	% Ileo-colonic		50.4	53.5	
	% Concomitant upper disease		5.4	4.2	.37
Behavior					.11
	% Inflammatory		53.4	48.1	
	% Stricturing		27.8	27.2	
	% Penetrating		18.9	24.6	
% Perianal disease			14.6	14.6	.99
Mean Harvey-Bradshaw Index[a]			3.6	4.4	<.01

XSL•FO
RenderX

		Linked (n=436)	Unlinked (n=1792)	*P* value
Ulcerative colitis				<.01
	% Proctitis	14.7	13.1	
	% Left-sided colitis	40.4	27.4	
	% Extensive colitis	44.9	59.5	
Mean Clinical Colitis Activity Index[a]		3.6	3.6	.52

[a]At the time of contact

Discussion

Principal Findings

This work represents a novel effort to link traditional cohorts of subjects with IBDs to an electronic cohort of IBD subjects. OSCCAR is the only US-based inception cohort of adult patients with IBD. SHARE is a large multicentered cohort of IBD subjects created to facilitate translational research. The accompanying biobank of specimens is an invaluable aspect of these cohorts. CCFA Partners is the largest US cohort of IBD subjects with a primary focus on patient-reported health data. The opportunity to leverage and enhance both cohorts through corecruitment and linkage of data assets is quite intuitive. The OSCCAR and SHARE cohorts are enriched through inclusion of an expanded array of patient-reported data. Similarly, CCFA Partners can provide a low-cost opportunity for the long-term follow-up of OSCCAR and SHARE participants in the event that funding for these cohorts ended. Conversely, for the portion of CCFA Partners participants who are coenrolled in OSCCAR or SHARE, the clinical data and biospecimens represent a valuable resource. Similar linkages between CCFA Partners and other clinical/translational cohorts will further expand the amount and types of data available to researchers.

As the findings reveal, it was a challenge to achieve linkage between the cohorts. Study staff worked diligently to use every contact as an opportunity to encourage subjects to participate in CCFA Partners. Although nearly 76% of the subjects who were eligible initially consented to enroll in CCFA Partners and link data, ultimately only 14% of those subjects who were eligible completed the linkage. The SHARE cohort had slightly more success with 24% of subjects linking to the CCFA Partners cohort. There are still a number of valuable lessons to be learned from this effort that both further inform the CCFA Partners study and may help inform future research endeavors involving record linkages.

CCFA Partners is an Internet-based cohort with most referrals coming from the CCFA rather than a population-based or randomly sampled cohort. The degree to which CCFA Partners is more broadly representative of the larger IBD population is largely unknown. However, systematic recruitment through the OSCCAR community-based cohort and SHARE multicentered cohort provided us the first opportunity to assess the generalizability of CCFA Partners. Education, race, gender, and disease activity for CD and location for UC appear to be factors impacting the decision to participate in CCFA Partners. This is important in understanding the generalizability of past and future findings learned through CCFA Partners. We also note that OSCCAR or SHARE participants who did not have access to the Internet were excluded from this linkage study. Therefore, additional demographic or clinical differences between the CCFA Partners cohort and the general IBD population may exist and will need to be evaluated in subsequent studies.

Data exploring factors contributing to participation in cohort studies of chronic diseases are scarce. Our findings indicate that education level was a very important driver of cohort linkage. In the OSCCAR study, 17% of participants who successfully linked possessed a doctoral degree compared with 3% of participants who did not complete the linkage. Furthermore, a higher percentage of participants who did not link had less than a bachelor's degree. One possible explanation for this difference is that those with a doctoral degree valued the importance of research endeavors and made it a priority to complete the linkage. Education level was not available for the SHARE cohort.

The SHARE experience revealed that gender and race may play a role in completing linkage to an Internet-based cohort, with a significantly higher percentage of women and white people completing the linkage. It is possible that those who linked have less aggressive disease than those who did not link. Among CD patients, although there were no differences in disease phenotypes, those who linked had a lower mean HBI. Among UC patients, those who linked had less-extensive disease, even though the mean SCCAI was the same between both groups.

To date, there is only one study examining factors in participation in cohort studies of IBDs. A Swiss IBD cohort reported that factors associated with nonresponse to an initial survey were younger age, male gender, and depression in CD only and longer disease duration in UC only. [31] These differences were not noted in our experience. HIV/AIDS is the disease process in which study participation has been most studied. Unemployment, female gender, higher education, and higher baseline CD4 counts predicted likelihood of participation in HIV cohort studies [22-34]. A study of childhood cancer survivors concluded that there were no substantive differences explaining nonparticipation in a lifetime cohort study [35], similar to our findings. Other groups described factors that contribute to participation in studies in general. A cohort study of employees at a large French utility company demonstrated that male sex, being married, having children, being a manager, and geography were predictors of participation [36]. However,

this study likely cannot be generalized to cohort studies of disease processes.

We speculate that recruiting participants into CCFA Partners at the time of their enrollment in OSCCAR (the time of their initial diagnosis) would have likely yielded more success in linkage. The OSCCAR cohort started enrollment 3 years before the initiation of the Partners cohort. It is possible that research fatigue was a factor explaining the poor linkage. It is also possible that the farther participants get from their diagnosis, the more their initial enthusiasm to participate in research wanes. SHARE participants were recruited at the time of the initial SHARE consent. Although similar proportions of subjects consented to enrollment, the SHARE experience demonstrated slightly more success with completing the linkage. Future investigators establishing new cohorts may want to anticipate potential linkages and incorporate them into the initial procedures at the beginning of the cohort. Indeed this is already being implemented with CCFA Partners recruitment and linkage incorporated into the design of the recently funded IBD Plexus Prospective Clinical Cohort.

The main strength of this study was the leveraging of the existing infrastructure of 3 well-developed cohorts. The primary limitation, aside from the disappointing recruitment and linkage results, is that we were not able to further explore reasons for noncompletion of CCFA Partners enrollment and linkage. Although interviews and surveys of OSCCAR and SHARE participants to inquire about reasons for nonparticipation may have been informative, we decided not to pursue this due to a concern that this might jeopardize continued participation in the cohorts themselves.

Conclusions

The linking of complementary cohorts may be a cost-effective strategy for expanding research infrastructure and enabling studies that might not be possible with either cohort alone. The regulatory, legal, and technical challenges in creating and implementing such linkages are not insurmountable; however, in this pilot study, we observed that recruitment itself was the substantial obstacle. Therefore, recruitment efforts for cohort linkages need to be balanced against the potential returns. Based on our experience, we would recommend a simultaneous, rather than sequential, strategy for corecruitment for subsequent attempts at cohort linkage. We believe that the lessons learned from this novel effort serve well to inform future linkages both in IBD and other chronic diseases.

Conflicts of Interest

Dr Kappelman is a consultant to GlaxoSmithKline and owns stock in GlaxoSmithKline. Suzanne Cook owns stock in GlaxoSmithKline.

References

1. Grodis L. Epidemiology. Philadelphia, PA: Elsevier; Jan 01, 2004.
2. Fone DL, Dunstan F, White J, Kelly M, Farewell D, John G, et al. Cohort profile: the Caerphilly health and social needs electronic cohort study (E-CATALyST). Int J Epidemiol 2013 Dec;42(6):1620-1628 [FREE Full text] [doi: 10.1093/ije/dys175] [Medline: 23132614]
3. Tuckett A, Hughes K, Gilmour J, Hegney D, Huntington A, Turner C. Caring in residential aged-care. Qualitative findings from an e-cohort sub-study. J Clin Nurs 2009 Sep;18(18):2604-2612. [doi: 10.1111/j.1365-2702.2008.02735.x] [Medline: 19538567]
4. Alfonsson S, Olsson E, Hursti T. Motivation and Treatment Credibility Predicts Dropout, Treatment Adherence, and Clinical Outcomes in an Internet-Based Cognitive Behavioral Relaxation Program: A Randomized Controlled Trial. J Med Internet Res 2016;18(3):e52 [FREE Full text] [doi: 10.2196/jmir.5352] [Medline: 26957354]
5. Kelstrup AM, Juillerat P, Korzenik J. The accuracy of self-reported medical history: a preliminary analysis of the promise of internet-based research in Inflammatory Bowel Diseases. J Crohns Colitis 2014 May;8(5):349-356. [doi: 10.1016/j.crohns.2013.09.012] [Medline: 24183653]
6. Roos LL, Menec V, Currie RJ. Policy analysis in an information-rich environment. Soc Sci Med 2004 Jun;58(11):2231-2241. [doi: 10.1016/j.socscimed.2003.08.008] [Medline: 15047080]
7. Jutte DP, Roos LL, Brownell MD. Administrative record linkage as a tool for public health research. Annu Rev Public Health 2011;32:91-108. [doi: 10.1146/annurev-publhealth-031210-100700] [Medline: 21219160]
8. Matta R, Doiron C, Leveridge MJ. The dramatic increase in social media in urology. J Urol 2014 Aug;192(2):494-498. [doi: 10.1016/j.juro.2014.02.043] [Medline: 24576656]
9. Ashley L, Jones H, Forman D, Newsham A, Brown J, Downing A, et al. Feasibility test of a UK-scalable electronic system for regular collection of patient-reported outcome measures and linkage with clinical cancer registry data: the electronic Patient-reported Outcomes from Cancer Survivors (ePOCS) system. BMC Med Inform Decis Mak 2011;11:66 [FREE Full text] [doi: 10.1186/1472-6947-11-66] [Medline: 22029686]

10. Sands BE, LeLeiko N, Shah SA, Bright R, Grabert S. OSCCAR: Ocean State Crohn's and Colitis Area Registry. Med Health R I 2009 Mar;92(3):82-5, 88. [Medline: 19385383]

11. Tinsley A, Macklin EA, Korzenik JR, Sands BE. Validation of the functional assessment of chronic illness therapy-fatigue (FACIT-F) in patients with inflammatory bowel disease. Aliment Pharmacol Ther 2011 Dec;34(11-12):1328-1336 [FREE Full text] [doi: 10.1111/j.1365-2036.2011.04871.x] [Medline: 21999576]

12. Pallis AG, Mouzas IA, Vlachonikolis IG. The inflammatory bowel disease questionnaire: a review of its national validation studies. Inflamm Bowel Dis 2004 May;10(3):261-269. [Medline: 15290922]

13. Otley A, Smith C, Nicholas D, Munk M, Avolio J, Sherman PM, et al. The IMPACT questionnaire: a valid measure of health-related quality of life in pediatric inflammatory bowel disease. J Pediatr Gastroenterol Nutr 2002 Oct;35(4):557-563. [Medline: 12394384]

14. Brazier JE, Harper R, Jones NM, O'Cathain A, Thomas KJ, Usherwood T, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. BMJ 1992 Jul 18;305(6846):160-164 [FREE Full text] [Medline: 1285753]

15. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. Med Care 2005 Mar;43(3):203-220. [Medline: 15725977]

16. Irvine EJ, Zhou Q, Thompson AK. The Short Inflammatory Bowel Disease Questionnaire: a quality of life instrument for community physicians managing inflammatory bowel disease. CCRPT Investigators. Canadian Crohn's Relapse Prevention Trial. Am J Gastroenterol 1996 Aug;91(8):1571-1578. [Medline: 8759664]

17. Kroenke K, Strine TW, Spitzer RL, Williams Janet B W, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1-3):163-173. [doi: 10.1016/j.jad.2008.06.026] [Medline: 18752852]

18. Cohen BL, Zoëga H, Shah SA, Leleiko N, Lidofsky S, Bright R, et al. Fatigue is highly associated with poor health-related quality of life, disability and depression in newly-diagnosed patients with inflammatory bowel disease, independent of disease activity. Aliment Pharmacol Ther 2014 Apr;39(8):811-822 [FREE Full text] [doi: 10.1111/apt.12659] [Medline: 24612278]

19. Harvey RF, Bradshaw JM. A simple index of Crohn's-disease activity. Lancet 1980 Mar 8;1(8167):514. [Medline: 6102236]

20. Walmsley RS, Ayres RC, Pounder RE, Allan RN. A simple clinical colitis activity index. Gut 1998 Jul;43(1):29-32 [FREE Full text] [Medline: 9771402]

21. Shen B, Achkar J, Connor JT, Ormsby AH, Remzi FH, Bevins CL, et al. Modified pouchitis disease activity index: a simplified approach to the diagnosis of pouchitis. Dis Colon Rectum 2003 Jun;46(6):748-753. [doi: 10.1097/01.DCR.0000070528.00563.D9] [Medline: 12794576]

22. Clara I, Lix LM, Walker JR, Graff LA, Miller N, Rogala L, et al. The Manitoba IBD Index: evidence for a new and simple indicator of IBD activity. Am J Gastroenterol 2009 Jul;104(7):1754-1763. [doi: 10.1038/ajg.2009.197] [Medline: 19455122]

23. Long MD, Kappelman MD, Martin CF, Lewis JD, Mayer L, Kinneer PM, et al. Development of an internet-based cohort of patients with inflammatory bowel diseases (CCFA Partners): methodology and initial results. Inflamm Bowel Dis 2012 Nov;18(11):2099-2106. [doi: 10.1002/ibd.22895] [Medline: 22287300]

24. Long MD, Kappelman MD, Martin CF, Chen W, Anton K, Sandler RS. Risk factors for depression in the elderly inflammatory bowel disease population. J Crohns Colitis 2014 Feb;8(2):113-119 [FREE Full text] [doi: 10.1016/j.crohns.2013.07.002] [Medline: 23932782]

25. Kappelman MD, Long MD, Martin C, DeWalt DA, Kinneer PM, Chen W, et al. Evaluation of the patient-reported outcomes measurement information system in a large cohort of patients with inflammatory bowel diseases. Clin Gastroenterol Hepatol 2014 Aug;12(8):1315-23.e2 [FREE Full text] [doi: 10.1016/j.cgh.2013.10.019] [Medline: 24183956]

26. Herfarth HH, Martin CF, Sandler RS, Kappelman MD, Long MD. Prevalence of a gluten-free diet and improvement of clinical symptoms in patients with inflammatory bowel diseases. Inflamm Bowel Dis 2014 Jul;20(7):1194-1197 [FREE Full text] [doi: 10.1097/MIB.0000000000000077] [Medline: 24865778]

27. Cohen AB, Lee D, Long MD, Kappelman MD, Martin CF, Sandler RS, et al. Dietary patterns and self-reported associations of diet with symptoms of inflammatory bowel disease. Dig Dis Sci 2013 May;58(5):1322-1328 [FREE Full text] [doi: 10.1007/s10620-012-2373-3] [Medline: 22923336]

28. Ananthakrishnan AN, Long MD, Martin CF, Sandler RS, Kappelman MD. Sleep disturbance and risk of active disease in patients with Crohn's disease and ulcerative colitis. Clin Gastroenterol Hepatol 2013 Aug;11(8):965-971 [FREE Full text] [doi: 10.1016/j.cgh.2013.01.021] [Medline: 23376797]

29. Jones PD, Kappelman MD, Martin CF, Chen W, Sandler RS, Long MD. Exercise decreases risk of future active disease in patients with inflammatory bowel disease in remission. Inflamm Bowel Dis 2015 May;21(5):1063-1071. [doi: 10.1097/MIB.0000000000000333] [Medline: 25723616]

30. Randell RL, Long MD, Cook SF, Wrennall Christina E D, Chen W, Martin CF, et al. Validation of an internet-based cohort of inflammatory bowel disease (CCFA partners). Inflamm Bowel Dis 2014 Mar;20(3):541-544 [FREE Full text] [doi: 10.1097/01.MIB.0000441348.32570.34] [Medline: 24451221]

31.   Multone E, Vader J, Mottet C, Schoepfer A, Fournier N, Burnand B, et al. Characteristics of non-responders to self-reported questionnaires in a large inflammatory bowel disease cohort study. Scand J Gastroenterol 2015;50(11):1348-1356. [doi: 10.3109/00365521.2015.1041150] [Medline: 25921677]

32.   Battegay M, Wirz M, Steuerwald MH, Egger M. Early participation in an HIV cohort study slows disease progression and improves survival. The Swiss HIV Cohort Study. J Intern Med 1998 Dec;244(6):479-487 [FREE Full text] [Medline: 9893101]

33.   Kim Y, Peragallo N, DeForge B. Predictors of participation in an HIV risk reduction intervention for socially deprived Latino women: a cross sectional cohort study. Int J Nurs Stud 2006 Jul;43(5):527-534. [doi: 10.1016/j.ijnurstu.2005.07.005] [Medline: 16153646]

34.   Thierfelder C, Weber R, Elzi L, Furrer H, Cavassini M, Calmy A, Swiss HIV Cohort Study Group. Participation, characteristics and retention rates of HIV-positive immigrants in the Swiss HIV Cohort Study. HIV Med 2012 Feb;13(2):118-126 [FREE Full text] [doi: 10.1111/j.1468-1293.2011.00949.x] [Medline: 22107170]

35.   Ojha RP, Oancea SC, Ness KK, Lanctot JQ, Srivastava DK, Robison LL, et al. Assessment of potential bias from non-participation in a dynamic clinical cohort of long-term childhood cancer survivors: results from the St. Jude Lifetime Cohort Study. Pediatr Blood Cancer 2013 May;60(5):856-864 [FREE Full text] [doi: 10.1002/pbc.24348] [Medline: 23024097]

36.   Goldberg M, Chastang JF, Leclerc A, Zins M, Bonenfant S, Bugel I, et al. Socioeconomic, demographic, occupational, and health factors associated with participation in a long-term epidemiologic survey: a prospective study of the French GAZEL cohort and its target population. Am J Epidemiol 2001 Aug 15;154(4):373-384 [FREE Full text] [Medline: 11495861]

Original Paper

Medical Registry Data Collection Efficiency: A Crossover Study Comparing Web-Based Electronic Data Capture and a Standard Spreadsheet

Pedro Vinícius Staziaki[1], MD; Phillip Kim[1]; Harshna V Vadvala[1], MD; Brian B Ghoshhajra[1], MD, MBA

Massachusetts General Hospital, Department of Radiology, Harvard Medical School, Boston, MA, United States

Corresponding Author:
Brian B Ghoshhajra, MD, MBA
Massachusetts General Hospital
Department of Radiology
Harvard Medical School
165 Cambridge St
Suite 400
Boston, MA, 02114
United States
Phone: 1 6177263745
Fax: 1 6177244152
Email: bghoshhajra@mgh.harvard.edu

Abstract

Background: Electronic medical records and electronic data capture (EDC) have changed data collection in clinical and translational research. However, spreadsheet programs, such as Microsoft Excel, are still used as data repository to record and organize patient data for research.

Objective: The objective of this study is to assess the efficiency of EDC as against a standard spreadsheet in regards to time to collect data and data accuracy, measured in number of errors after adjudication.

Methods: This was a crossover study comparing the time to collect data in minutes between EDC and a spreadsheet. The EDC tool used was Research Electronic Data Capture (REDCap), whereas the spreadsheet was Microsoft Excel. The data collected was part of a registry of patients who underwent coronary computed tomography angiography in the emergency setting. Two data collectors with the same experience went over the same patients and collected relevant data on a case report form identical to the one used in our Emergency Department (ED) registry. Data collection tool was switched after the patient that represented half the cohort. For this, the patient cohort was exactly 30 days of our ED coronary Computed Tomography Angiography registry and the point of crossover was determined beforehand to be 15 days. We measured the number of patients admitted, and time to collect data. Accuracy was defined as absence of blank fields and errors, and was assessed by comparing data between data collectors and counting every time the data differed. Statistical analysis was made using paired t-test.

Results: The study included 61 patients (122 observations) and 55 variables. The crossover occurred after the 30th patient. Mean time to collect data using EDC in minutes was 6.2±2.3, whereas using Excel was 8.0±2.0 ($P <.001$), a difference of 1.8 minutes between both means (22%). The cohort was evenly distributed with 3 admissions in the first half of the crossover and 4 in the second half. We saw 2 (<0.1%) continuous variable typos in the spreadsheet that a single data collector made. There were no blank fields. The data collection tools showed no differences in accuracy of data on comparison.

Conclusions: Data collection for our registry with an EDC tool was faster than using a spreadsheet, which in turn allowed more efficient follow-up of cases.

(*J Med Internet Res 2016;18(6):e141*)   doi:10.2196/jmir.5576

XSL•FO

RenderX

Introduction

Electronic medical records and electronic data capture (EDC) have changed data collection in clinical and translational research [1]. Electronic forms reduce inaccurate data entry and study costs because the data are entered directly into an electronic form on a computer [2]. However, spreadsheet programs are still used as data repository to record and organize patient data for research. This method of data storage is not only limited in the organization and quality of data but also increases the likelihood of incorrect data entry [3].

It is known that EDC reduces cost and time when compared with paper-based data collection [4-6]. However, there is little research about how EDC solutions are better compare with spreadsheets. Furthermore, most of the literature about EDC is descriptive, focusing only on the technology, the methods, or the experience [7].

The objective of this study was to assess the efficiency of an EDC solution compared with a standard spreadsheet regarding time to collect data and data accuracy, measured in number of errors after adjudication. We hypothesized that EDC reduces the time of data collection without compromising accuracy, as compared with a standard spreadsheet.

Methods

This was a single-institution crossover study comparing the time to collect data in minutes between an EDC tool and a spreadsheet. This study was approved by the Institutional Review Board and was Health Insurance Portability and Accountability Act (HIPAA) compliant.

Study Design

Two data collectors ("1" and "2") went over the same patients and collected relevant clinical and imaging data, switching data collection tool after the patient that represented half the cohort (Figure 1). Both data collectors observed each patient, one collecting data on EDC and other on a spreadsheet.

We designed this study to simulate the actual registry data collection environment. For this, the patient cohort was exactly 30 days of our Emergency Department (ED) coronary Computed Tomography Angiography registry and the point of crossover was determined beforehand to be 15 days. The case report form (CRF) for this study was the same as used in our ED registry (Figure 2).

Anticipating that certain patients would be admitted to the hospital and contain more data to be collected, we also looked at how many of those patients were admitted, in order to know if they were evenly distributed between each half of the crossover.

The EDC tool used was Research Electronic Data Capture (REDCap) [8] and the spreadsheet application was Excel (Microsoft Corporation, Redmond, Washington).

Each data collectors had 5 months' experience in registry data collection and used the same versions of REDCap and Excel and an electronic medical record system (QPID). Both users worked on the same computer systems having the same Internet speed. The CRFs on each data collection tool had identical variables, which comprised dichotomous variables, categorical variables, and continuous variables.

The time to collect data was recorded in a separate spreadsheet (Figure 3). Both data collectors recorded time identically irrespective of the tool (spreadsheet or EDC) used by them for the registry data collection. This spreadsheet was different from the data collection spreadsheet that was to be compared with EDC.

Figure 1. Crossover design.

XSL•FO
RenderX
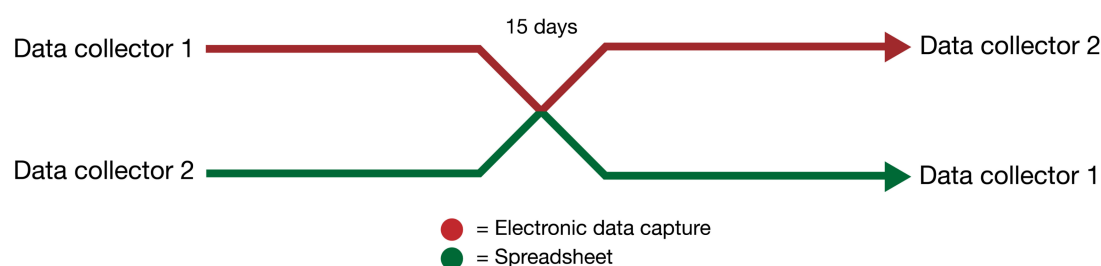
Figure 2. A sample of our ED registry case report form (CRF).

Figure 3. Sample of the first rows of time data for data collector "1", for each data collection tool. The full cohort contained 61 rows. Column A identifies the data collector ("1" or "2"). Column B identifies the study subject (each registry record) with their accession number (ACC), here anonymized. Column C indicates the tool used to collect the registry data. Column D indicates the time-stamp of the start of data collection, whereas column E indicates the end. Column F contains the calculated time of data collection in minutes (column E minus column D). Admission was coded in column G as 1 for presence and 0 for absence. EDC = Electronic data capture.

	A	B	C	D	E	F	G
1	Data collector	ACC	Tool	Begin	End	Time	Admission
2	1	1	EDC	3:24	3:32	8	0
3	1	2	EDC	3:35	3:48	13	1
4	1	3	EDC	3:49	3:56	7	0
5	1	4	EDC	3:57	4:06	9	0
6	1	5	EDC	4:07	4:16	9	0
7	1	6	EDC	4:17	4:25	8	0
8	1	7	EDC	4:25	4:32	7	0
			...				
31	1	30	EDC	11:19	11:25	6	0
32	1	31	Spreadsheet	12:18	12:30	12	0
33	1	32	Spreadsheet	12:54	13:05	11	0
34	1	33	Spreadsheet	1:06	1:17	11	0
35	1	34	Spreadsheet	1:21	1:29	8	0
36	1	35	Spreadsheet	1:30	1:37	7	0
37	1	36	Spreadsheet	11:07	11:20	13	1
38	1	37	Spreadsheet	11:22	11:34	12	1
			...				
62	1	61	Spreadsheet	2:24	2:32	8	0

Time and Accuracy Data

In order to collect the time data, we manually typed the time stamp of the beginning of data collection in one column and the end of data collection in another column in the time spreadsheet. Time was calculated in minutes by subtracting the beginning time-stamp from the end time-stamp and it was recorded in the next column.

Accuracy was assessed by comparing three indicators between both data collection tools: the number of blank fields, the number of discrepant fields, and the type of discrepancies. For the type of discrepancy, we looked at every pair of record, comparing a record from one tool to the same record from the other tool.

The discrepancies were categorized into two groups: different content and same content errors, such as typos. Discrepancies that represented different content were adjudicated by a senior radiologist to select which record in each pair is deemed the wrong data entry.

Statistical Analysis

Statistical analysis was made using paired t-test. Every patient was tested twice, as each was collected once on REDCap and once on spreadsheet.

Results

The study included 61 patients (122 observations) and 55 variables. The crossover occurred after the 30th patient. Mean time to collect data using EDC in minutes was 6.2±2.3, whereas using a spreadsheet was 8.0±2.0 ($P <.001$), resulting in a reduction of 1.8 out of 8 minutes (22%). The cohort was evenly distributed, with 3 admissions in the first half of the crossover and 4 in the second half.

In all, 6710 entries of the registry were collected (61 patients × 55 variables, 2 collectors). We saw 2 continuous variable typos out of 6710 (<0.1%) that a single data collector made in Excel. There were no blank fields and no discrepancies.

Discussion

The main finding of this study was that less time is required to collect data to an EDC than to a spreadsheet. Prior literature

XSL•FO
RenderX

has compared EDC with conventional paper capture methods and it is mainly descriptive. This study compared objectively the time to collect data between a Web-based rigid form and a standard spreadsheet, and confirmed that EDC using REDCap can be more time effective. We chose to compare EDC to spreadsheets since we have found that in the era of electronic medical records, efficiency can be gained by using only EDC, and the final form of data delivered for research analysis is usually always electronic.

Regarding the time to collect a single data endpoint, a small difference in time can add up to a significant difference in the long term. It took 6.3 h to collect the data in REDCap compared with 8.1 hin Excel, a difference of 1.8 h. In our clinical registry of over 1000 ED admissions, this means that by collecting all data via this EDC solution we would spend only 103 h (6.2 min $\times$ 1000 observations) as opposed to 133 (8.0 $\times$ 1000), saving more than 3 workdays of data collection.

Concerning accuracy, there were no discrepancies between the two data collection tools. The number of errors was too small compared with the number of observations collected. Due to this, we did not perform a statistical analysis of the number of errors in data entry. In addition, since single data collector made typos in the spreadsheet, we did not see differences in data collection that could be attributed to a specific tool in our study.

Setting up ranges and automatic calculations can prevent these errors. Range checks make sure the collector does not insert a typo that would give a value in continuous variables that would not make sense [9]. While these can be set up in both Excel and REDCap, the latter can provide a better interface making it is easier to be done.

Many EDC tools have been analyzed [8,10-12]. The major advantage of electronic CRFs over spreadsheets is that the former can be designed to present only certain acceptable choices for an item or to check the syntax and range of data that are entered [9]. This reduces the likelihood of data entry errors.

REDCap was developed at Vanderbilt University's Institute for Clinical and Translational Research for building and managing online databases [8]. REDCap is an NIH-sponsored, HIPAA-compliant, noncommercial, and secure EDC solution. It supports retrospective and prospective studies, as well as multicenter clinical trials. It has an intuitive user-friendly interface for data entry, allowing researchers to create secure online forms with very large numbers and several types of variables and does not require any technical skillset to implement. By organizing the variables into forms, it also permits the user to save the progress between each form, avoiding the dramatic trouble of losing data by not saving the progress. These forms also allow for mid-study modifications without affecting previously collected data.

Data collections in standard spreadsheets can be easily imported to REDCap, and then data can be exported into most major statistical software packages, such as Stata (StataCorp, College Station, TX, USA), SAS (SAS Institute, Cary, NC, USA), R (R Foundation for Statistical Computing, Vienna, Austria), and SPSS (IBM Corporation, Armonk, New York), as well as comma-delimited files. As it is a Web-based tool, it is compatible with all operating systems [13] and requires no installation of software [14].

Irrespective of the data collection tool used, the data is often exported to a comma-delimited file that can be read as a spreadsheet. Then, this file can be imported into statistical software packages. Moreover, spreadsheets are the common file format through which researchers and statisticians exchange the data.

However, spreadsheets require the data collector to abide by certain practices regarding how data are organized and formatted within the spreadsheet [15], such as putting variable names in a single row and avoiding the use of special characters. Also, spreadsheets for data collection restricts features such as colored text, cell shading, commas, merging cells, comments, or mixing data types in a single column. This makes the data collection more time-consuming and error-prone.

This study has limitations. A limitation of our study is that it is a small dataset.

In addition, the times to collect data reported are inherent to our registry, and would be different in different research studies. Therefore, it is difficult to extrapolate our results to other research projects. Nevertheless, the use of a crossover design ensured the data was controlled, and this method accounted for differences in speed inherent to each collector.

In conclusion, data collection for our registry with EDC was faster than using a spreadsheet, allowing more efficient follow-up of cases.

Conflicts of Interest

None declared.

References

1. Welker JA. Implementation of electronic data capture systems: barriers and solutions. Contemp Clin Trials 2007 May;28(3):329-336. [doi: 10.1016/j.cct.2007.01.001] [Medline: 17287151]
2. Pavlović I, Kern T, Miklavcic D. Comparison of paper-based and electronic data collection process in clinical trials: costs simulation study. Contemp Clin Trials 2009 Jul;30(4):300-316. [doi: 10.1016/j.cct.2009.03.008] [Medline: 19345286]
3. Lee H, Chapiro J, Schernthaner R, Duran R, Wang Z, Gorodetski B, et al. How I do it: a practical database management system to assist clinical research teams with data collection, organization, and reporting. Acad Radiol 2015 Apr;22(4):527-533. [doi: 10.1016/j.acra.2014.12.002] [Medline: 25641319]

XSL•FO

RenderX

4.   Babre D. Electronic data capture - Narrowing the gap between clinical and data management. Perspect Clin Res 2011 Jan;2(1):1-3 [FREE Full text] [doi: 10.4103/2229-3485.76282] [Medline: 21584174]

5.   Walther B, Hossin S, Townend J, Abernethy N, Parker D, Jeffries D. Comparison of electronic data capture (EDC) with the standard data capture method for clinical trial data. PLoS One 2011;6(9):e25348 [FREE Full text] [doi: 10.1371/journal.pone.0025348] [Medline: 21966505]

6.   Franklin JD, Guidry A, Brinkley JF. A partnership approach for Electronic Data Capture in small-scale clinical trials. J Biomed Inform 2011 Dec;44 Suppl 1:S103-S108 [FREE Full text] [doi: 10.1016/j.jbi.2011.05.008] [Medline: 21651992]

7.   Pawellek I, Richardsen T, Oberle D, Grote V, Koletzko B. Use of electronic data capture in a clinical trial on infant feeding. Eur J Clin Nutr 2012 Dec;66(12):1342-1343. [doi: 10.1038/ejcn.2012.141] [Medline: 23211655]

8.   Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]

9.   Kupzyk KA, Cohen MZ. Data validation and other strategies for data entry. West J Nurs Res 2015 Apr;37(4):546-556. [doi: 10.1177/0193945914532550] [Medline: 24782433]

10. Rangel SJ, Narasimhan B, Geraghty N, Moss RL. Development of an internet-based protocol to facilitate randomized clinical trials in pediatric surgery. J Pediatr Surg 2002 Jul;37(7):990-4; discussion 990. [Medline: 12077756]

11. Marks R, Bristol H, Conlon M, Pepine CJ. Enhancing clinical trials on the internet: lessons from INVEST. Clin Cardiol 2001 Nov;24(11 Suppl):V17-V23 [FREE Full text] [Medline: 11712772]

12. Lallas CD, Preminger GM, Pearle MS, Leveillee RJ, Lingeman JE, Schwope JP, et al. Internet based multi-institutional clinical research: a convenient and secure option. J Urol 2004 May;171(5):1880-1885. [doi: 10.1097/01.ju.0000120221.39184.3c] [Medline: 15076298]

13. Tuti T, Bitok M, Paton C, Makone B, Malla L, Muinga N, et al. Innovating to enhance clinical data management using non-commercial and open source solutions across a multi-center network supporting inpatient pediatric care and research in Kenya. J Am Med Inform Assoc 2016 Jan;23(1):184-192 [FREE Full text] [doi: 10.1093/jamia/ocv028] [Medline: 26063746]

14. Paul J, Seib R, Prescott T. The Internet and clinical trials: background, online resources, examples and issues. J Med Internet Res 2005;7(1):e5 [FREE Full text] [doi: 10.2196/jmir.7.1.e5] [Medline: 15829477]

15. Juluru K, Eng J. Use of Spreadsheets for Research Data Collection and Preparation: A Primer. Acad Radiol 2015 Dec;22(12):1592-1599. [doi: 10.1016/j.acra.2015.08.024] [Medline: 26454810]

Abbreviations

CRF: case report form
ED: emergency department
EDC: electronic data capture

<u>Short Paper</u>

Web-Based Assessment of Mental Well-Being in Early Adolescence: A Reliability Study

Christoph Hamann[1], MD; Frauke Schultze-Lutter[1], PhD; Leila Tarokh[1,2,3], PhD

[1]University Hospital of Child and Adolescent Psychiatry and Psychotherapy, University of Bern, Bern, Switzerland

[2]Institute of Pharmacology and Toxicology, University of Zurich, Zurich, Switzerland

[3]Psychiatry and Human Behavior, The Alpert Medical School of Brown University, Providence, RI, United States

Corresponding Author:
Leila Tarokh, PhD
University Hospital of Child and Adolescent Psychiatry and Psychotherapy
University of Bern
Bolligenstr. 111, Haus A
Bern, 3000
Switzerland
Phone: 41 31 932 8554
Fax: 41 31 932 8569
Email: leila_tarokh@brown.edu

Abstract

Background: The ever-increasing use of the Internet among adolescents represents an emerging opportunity for researchers to gain access to larger samples, which can be queried over several years longitudinally. Among adolescents, young adolescents (ages 11 to 13 years) are of particular interest to clinicians as this is a transitional stage, during which depressive and anxiety symptoms often emerge. However, it remains unclear whether these youngest adolescents can accurately answer questions about their mental well-being using a Web-based platform.

Objective: The aim of the study was to examine the accuracy of responses obtained from Web-based questionnaires by comparing Web-based with paper-and-pencil versions of depression and anxiety questionnaires.

Methods: The primary outcome was the score on the depression and anxiety questionnaires under two conditions: (1) paper-and-pencil and (2) Web-based versions. Twenty-eight adolescents (aged 11-13 years, mean age 12.78 years and SD 0.78; 18 females, 64%) were randomly assigned to complete either the paper-and-pencil or the Web-based questionnaire first. Intraclass correlation coefficients (ICCs) were calculated to measure intrarater reliability. Intraclass correlation coefficients were calculated separately for depression (Children's Depression Inventory, CDI) and anxiety (Spence Children's Anxiety Scale, SCAS) questionnaires.

Results: On average, it took participants 17 minutes (SD 6) to answer 116 questions online. Intraclass correlation coefficient analysis revealed high intrarater reliability when comparing Web-based with paper-and-pencil responses for both CDI (ICC=.88; $P<.001$) and the SCAS (ICC=.95; $P<.001$). According to published criteria, both of these values are in the "almost perfect" category indicating the highest degree of reliability.

Conclusions: The results of the study show an excellent reliability of Web-based assessment in 11- to 13-year-old children as compared with the standard paper-pencil assessment. Furthermore, we found that Web-based assessments with young adolescents are highly feasible, with all enrolled participants completing the Web-based form. As early adolescence is a time of remarkable social and behavioral changes, these findings open up new avenues for researchers from diverse fields who are interested in studying large samples of young adolescents over time.

KEYWORDS

early adolescence; online assessment; reliability

XSL•FO
RenderX

Introduction

Adolescent development is long, complex, and highly individual (eg, [1-4]). In order to capture the dynamic process that is adolescent development, large sample sizes or longitudinal assessment is often required [5]. Access to large samples of adolescents over extended periods can be challenging because of factors such as changing of schools, moving, or unstable family environments [6]. Thus, data collection is often cumbersome, expensive, and can result in dropouts that may influence results [7,8].

One way to circumvent some of these difficulties is through the administration of Web-based questionnaires. Adolescents today have regular access to the Internet either at school or at home. For example, according to the Pew Research Center, 95% of teenagers (12-17 years) have Internet access in the United States [9], and on average a European teenager uses online media for 66 minutes per day [10]. The ever-increasing use of the Internet among adolescents represents an emerging opportunity for researchers to gain access to large samples, which can be queried over several years longitudinally. Subjects can complete such questionnaires wherever and whenever it suits them using different media, for example, tablets, laptops, or mobile phones, and researchers can monitor data quality instantaneously. Furthermore, the Internet offers diverse options for communication between participants and researchers, which can be used to bolster participation and to minimize dropouts [8].

Previous studies have shown that adolescents (13-20 years) [11,12] can accurately and reliably fill out Web-based questionnaires about their mental and physical well-being. However, few studies have addressed the reliability of Web-based questionnaires in young adolescents between the ages of 11 and 13 years in a naturalistic setting (ie, at home, unmonitored). This age range is of particular interest to clinicians and researchers, because it is a transitional stage [13]. This transition from childhood to adolescence—accompanied by increased independence, a new school environment, the onset of puberty, and shifting peer relationships—can be highly stressful and may lead to the emergence of depressive and anxiety symptoms. The influence of stressors on psychopathology in youth is under intensive investigation and conclusions are still difficult to draw [14]. Understanding the etiology of psychiatric disorders is critical to early intervention and most disorders have their onset during adolescence [15-18]. Therefore, surveying adolescents early on in development and following them longitudinally will further our understanding of adolescent development.

Thus, the aim of this study was to examine whether young adolescents can accurately fill out Web-based questionnaires about depression and anxiety, the two most common psychiatric disorders among adolescents (lifetime prevalence of 25% of anxiety disorders and 13% of mood disorders in 13- to 18-year-olds [18]). We accomplish this by comparing Web-based with standard paper-and-pencil versions of depression and anxiety questionnaires.

Methods

Sample and Design

Twenty-eight children between the ages of 11 and 13 years (mean 12.78 years and SD 0.78; 18 females, 64%) participated in this study. The study was briefly introduced to 2 classes at a Swiss secondary school. Students who wished to participate sent back their contact information in addition to a consent form signed by their parents or legal guardian and themselves. Once recruited into the study, participants were randomly assigned to complete either the paper-and-pencil version or the Web-based version first in a randomized crossover design. In the paper-and-pencil condition, participants received the questionnaires by mail and a self-addressed stamped envelope was provided for returning the questionnaires. In the Web-based condition, participants received an identification number, a password, and a link to the survey website. Parents were instructed to leave the child alone to fill out the Web-based and paper-and-pencil forms. After completing the first condition, participants received the questionnaires of the alternate condition so that the second assessment was not more than 2 weeks from the first. This rather short lag time was chosen to limit the influence of changes in mental state on the completion of the state-sensitive questionnaires. Participants received compensation in the form of gift vouchers for taking part in the study, which was approved by the Ethics Committee in Bern, Switzerland.

Questionnaires

The German version (Depressions-Inventar für Kinder und Jugendliche, DIKJ) of the Children's Depression Inventory (CDI; [19]) was used to measure depressive symptoms. This scale is a well-established self-report measure of depressive symptoms appropriate for children between the ages of 7 and 17 years. This scale consists of 26 items, which are each scored from 0 to 2, and thus the total score of this scale yields values ranging between 0 and 52.

The German version of the Spence Children's Anxiety Scale (SCAS) was used to measure anxiety symptoms. This scale was designed to assess anxiety symptoms in individuals between the ages of 8 and 15 years and consists of 38 items, which are scored on a 4-point scale—thus scores range between 0 and 114. This scale also permits calculation of subscores, allowing for the evaluation of anxiety across specific domains [20-22].

Statistical Analysis

Statistical analysis was conducted with SPSS version 23.0.0.0. Intraclass correlation coefficients (ICCs) were used to examine the degree of correspondence between paper and Web-based versions. Intraclass correlation coefficient values range between 0 and 1 and are conventionally categorized as follows: 0-.2 poor, .2-.4 fair, .4-.6 moderate, .6-.8 substantial, and .8-1.0 almost perfect [23]. The Web-based version was considered reliable when an ICC value equal to or greater than .8 (minimum of substantial agreement) was obtained. Additionally, Kendall tau correlations were used to measure the equivalence between the paper-and-pencil and Web-based versions, and Wilcoxon tests

were performed to test for statistically significant differences between the two conditions.

Results

On average, it took 17 minutes (SD 6) to fill out the Web-based questionnaires, which consisted of the SCAS, CDI, and an additional 52 questions about sleep behavior and quality resulting in a total of 116 questions. Because paper-and-pencil questionnaires were filled out at home, no data on time taken to complete the forms were available. The evaluation of the paper-and-pencil and Web-based versions resulted in an overall ICC of .88 (*P*<.001) for the CDI and an ICC of .95 (*P*<.001) for the SCAS (Table 1). According to the criteria of Landis and Koch [23], ICC values higher than .8 fall into the category of "almost perfect" indicating the highest degree of reliability. Kendall tau correlation was also significant with correlation coefficients (*r*) of .58 (*P*<.001) for the CDI and .72 (*P*<.001) for the SCAS. Additionally, Wilcoxon tests comparing paper-and-pencil with the Web-based condition showed no significant difference (Table 1), with the exception of the subscore for Physical Injury Fear of the SCAS.

Table 1. The degree of correspondence between the paper-and-pencil and Web-based versions of depression and anxiety questionnaires filled out by 11- to 13-year-old children.

Questionnaire	Paper-and-pencil, mean (SD)	Web-based, mean (SD)	ICC[a] (*P*)	*Z*[b] (*P*)	τ[c] (*P*)
Total SCAS[d]	19.46 (12.64)	21.11 (13.41)	.95 (<.001)	−1.36 (.173)	.72 (<.001)
Separation Anxiety	2.43 (2.77)	2.75 (2.99)	.92 (<.001)	−0.82 (.410)	.75 (<.001)
Social Phobia	4.68 (2.94)	4.61 (3.01)	.91 (<.001)	−0.43 (.666)	.66 (<.001)
Obsessive Compulsivness	4.07 (3.03)	4.07 (2.98)	.92 (<.001)	−1.50 (.881)	.71 (<.001)
Panic Agoraphobia	2.26 (2.80)	1.78 (2.74)	.90 (<.001)	−1.40 (.163)	.49 (.002)
Physical Injury Fear	2.32 (2.18)	3.04 (2.78)	.90 (<.001)	−3.06 (.002)	.86 (<.001)
Generalized Anxiety	4.36 (2.36)	4.82 (2.61)	.86 (<.001)	−1.45 (.148)	.60 (<.001)
Total CDI[e] (DIKJ[f])	6.61 (4.46)	7.39 (4.41)	.88 (<.001)	−1.09 (.277)	.58 (<.001)

[a]ICC: intraclass correlation coefficient.

[b]Z: Wilcoxon-Mann-Whitney test.

[c]τ: Kendall tau-b.

[d]SCAS: Spence Children's Anxiety Scale.

[e]CDI: Children's Depression Inventory.

[f]DIKJ: Depressions-Inventar für Kinder und Jugendliche.

Discussion

In this randomized study, we show that Internet-based data collection of mental health questionnaires is feasible and reliable in early adolescence outside a highly supervised environment. A number of studies have reported feasibility and acceptance of Web-based questionnaires evaluating health among children [24-27], teenagers [11,12,26-29], and adults [30-40]; however, very few studies have examined the reliability of Web-based questionnaires under unsupervised conditions in a sample of young adolescents. For example, Mangunkusumo et al [25] found comparable responses on an Internet versus paper mode of health questionnaires in an elementary school cohort (ages 10-12 years), using a randomized within-subject design. However, the children in this study were intensively supervised by their teachers and changing modalities happened during the same school lesson with only a 5-minute break in between, thus making a memory bias in favor of high correspondence likely. Our within-subject results show a similar outcome in an unsupervised setting and with a longer lag time between assessments, where young adolescents filled out questionnaires during their free time without close supervision—an important modification for the feasibility of longitudinal study designs.

Furthermore, our results not only show high ICC values in the overall sum scores of CDI and SCAS, but they also show high ICC and correlation scores in the subscores of the SCAS. Only the Physical Injury Fear subscale differed between the Web-based and the paper-and-pencil groups and may be due to the small number of items that are included in each subscore.

Additionally, the adolescents in our study were able to answer a relatively large number of items (total of 116 items) in a short time, which shows the feasibility of collecting large datasets. Furthermore, despite our modest sample size, we were able to demonstrate high ICC values, which indicate high intrarater reliability in comparison to interrater variability. Nevertheless, the small sample size is a limitation of our study.

In summary, our study suggests that Internet-based data collection in young adolescents in the field of mental health and beyond is feasible and reliable. Consequently, Internet-based questionnaires can be implemented in larger longitudinal studies in the future in order to develop a better understanding of adolescent behavior and follow the emergence of psychiatric disorders, which can lead to prevention and better treatment.

Conflicts of Interest

None declared.

References

1. Andersen SL. Trajectories of brain development: point of vulnerability or window of opportunity? Neurosci Biobehav Rev 2003;27(1-2):3-18. [Medline: 12732219]

2. Giedd JN. The teen brain: insights from neuroimaging. J Adolesc Health 2008 Apr;42(4):335-343. [doi: 10.1016/j.jadohealth.2008.01.007] [Medline: 18346658]

3. Giedd JN, Rapoport JL. Structural MRI of pediatric brain development: what have we learned and where are we going? Neuron 2010 Sep 9;67(5):728-734 [FREE Full text] [doi: 10.1016/j.neuron.2010.08.040] [Medline: 20826305]

4. Berk LE. Development Through the Lifespan. 5th edition. Boston, MA: Pearson Education, Inc; 2009.

5. Farrington DP. Longitudinal research strategies: advantages, problems, and prospects. J Am Acad Child Adolesc Psychiatry 1991 May;30(3):369-374. [doi: 10.1097/00004583-199105000-00003] [Medline: 2055872]

6. Howe LD, Tilling K, Galobardes B, Lawlor DA. Loss to follow-up in cohort studies: bias in estimates of socioeconomic inequalities. Epidemiology 2013 Jan;24(1):1-9. [doi: 10.1097/EDE.0b013e31827623b1] [Medline: 23211345]

7. Gustavson K, Borren I. Bias in the study of prediction of change: a Monte Carlo simulation study of the effects of selective attrition and inappropriate modeling of regression toward the mean. BMC Med Res Methodol 2014;14:133 [FREE Full text] [doi: 10.1186/1471-2288-14-133] [Medline: 25519494]

8. David MC, Alati R, Ware RS, Kinner SA. Attrition in a longitudinal study with hard-to-reach participants was reduced by ongoing contact. J Clin Epidemiol 2013 May;66(5):575-581. [doi: 10.1016/j.jclinepi.2012.12.002] [Medline: 23384589]

9. Pew Research Center. 2013. Internet User Demographics URL: http://www.pewinternet.org/data-trend/teens/internet-user-demographics/ [accessed 2016-03-04] [WebCite Cache ID 6fl57y5EF]

10. Hermida M, Signer S. EU kids online. 2013. Willkommen bei EU Kids Online: Schweiz URL: http://www.eukidsonline.ch/ [accessed 2016-03-05] [WebCite Cache ID 6fl57OMvJ]

11. Wang Y, Lee C, Lew-Ting C, Hsiao CK, Chen D, Chen WJ. Survey of substance use among high school students in Taipei: web-based questionnaire versus paper-and-pencil questionnaire. J Adolesc Health 2005 Oct;37(4):289-295. [doi: 10.1016/j.jadohealth.2005.03.017] [Medline: 16182139]

12. Ramo DE, Liu H, Prochaska JJ. Reliability and validity of young adults' anonymous online reports of marijuana use and thoughts about use. Psychol Addict Behav 2012 Dec;26(4):801-811 [FREE Full text] [doi: 10.1037/a0026201] [Medline: 22082344]

13. Grills-Taquechel AE, Norton P, Ollendick TH. A longitudinal examination of factors predicting anxiety during the transition to middle school. Anxiety Stress Coping 2010 Oct;23(5):493-513 [FREE Full text] [doi: 10.1080/10615800903494127] [Medline: 20711893]

14. Kushner SC. A Review of the Direct and Interactive Effects of Life Stressors and Dispositional Traits on Youth Psychopathology. Child Psychiatry Hum Dev 2015 Oct;46(5):810-819. [doi: 10.1007/s10578-014-0523-x] [Medline: 25414134]

15. Thapar A, Collishaw S, Pine DS, Thapar AK. Depression in adolescence. Lancet 2012 Mar 17;379(9820):1056-1067 [FREE Full text] [doi: 10.1016/S0140-6736(11)60871-4] [Medline: 22305766]

16. Cummings CM, Caporino NE, Kendall PC. Comorbidity of anxiety and depression in children and adolescents: 20 years after. Psychol Bull 2014 May;140(3):816-845 [FREE Full text] [doi: 10.1037/a0034733] [Medline: 24219155]

17. 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. [doi: 10.1001/archpsyc.62.6.593] [Medline: 15939837]

18. Merikangas KR, He J, Burstein M, Swanson SA, Avenevoli S, Cui L, et al. Lifetime prevalence of mental disorders in U.S. adolescents: results from the National Comorbidity Survey Replication--Adolescent Supplement (NCS-A). J Am Acad Child Adolesc Psychiatry 2010 Oct;49(10):980-989 [FREE Full text] [doi: 10.1016/j.jaac.2010.05.017] [Medline: 20855043]

19. Kovacs M. The Children's Depression, Inventory (CDI). Psychopharmacol Bull 1985;21(4):995-998. [Medline: 4089116]

20. Spence SH. Structure of anxiety symptoms among children: a confirmatory factor-analytic study. J Abnorm Psychol 1997 May;106(2):280-297. [Medline: 9131848]

21. Spence SH. A measure of anxiety symptoms among children. Behav Res Ther 1998 May;36(5):545-566. [Medline: 9648330]

22. Spence SH, Barrett PM, Turner CM. Psychometric properties of the Spence Children's Anxiety Scale with young adolescents. J Anxiety Disord 2003;17(6):605-625. [Medline: 14624814]

23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]

24. McCabe SE, Boyd CJ, Young A, Crawford S, Pope D. Mode effects for collecting alcohol and tobacco data among 3rd and 4th grade students: a randomized pilot study of Web-form versus paper-form surveys. Addict Behav 2005 May;30(4):663-671. [doi: 10.1016/j.addbeh.2004.08.012] [Medline: 15833572]

25. Mangunkusumo RT, Duisterhout JS, de GN, Maarsingh EJ, de Koning Harry J, Raat H. Internet versus paper mode of health and health behavior questionnaires in elementary schools: asthma and fruit as examples. J Sch Health 2006 Feb;76(2):80-86. [doi: 10.1111/j.1746-1561.2006.00072.x] [Medline: 16466471]

26. Goodman R. Online child and adolescent mental health surveys can be good enough. Soc Psychiatry Psychiatr Epidemiol 2013 Aug;48(8):1317-1325. [doi: 10.1007/s00127-013-0658-2] [Medline: 23377755]

27. Hohwü L, Lyshol H, Gissler M, Jonsson SH, Petzold M, Obel C. Web-based versus traditional paper questionnaires: a mixed-mode survey with a Nordic perspective. J Med Internet Res 2013;15(8):e173 [FREE Full text] [doi: 10.2196/jmir.2595] [Medline: 23978658]

28. Raat H, Mangunkusumo RT, Landgraf JM, Kloek G, Brug J. Feasibility, reliability, and validity of adolescent health status measurement by the Child Health Questionnaire Child Form (CHQ-CF): internet administration compared with the standard paper version. Qual Life Res 2007 May;16(4):675-685 [FREE Full text] [doi: 10.1007/s11136-006-9157-1] [Medline: 17286197]

29. Raghupathy S, Hahn-Smith S. Reliability of the high school Youth Risk Behavior Survey when administered online. Int Q Community Health Educ 2011 Sep;32(2):135-148. [doi: 10.2190/IQ.32.2.d] [Medline: 23000460]

30. Coles M, Cook L, Blake T. Assessing obsessive compulsive symptoms and cognitions on the internet: evidence for the comparability of paper and Internet administration. Behav Res Ther 2007 Sep;45(9):2232-2240. [doi: 10.1016/j.brat.2006.12.009] [Medline: 17306222]

31. Holländare F, Andersson G, Engström I. A comparison of psychometric properties between internet and paper versions of two depression instruments (BDI-II and MADRS-S) administered to clinic patients. J Med Internet Res 2010;12(5):e49 [FREE Full text] [doi: 10.2196/jmir.1392] [Medline: 21169165]

32. Khazaal Y, Chatton A, Monney G, Nallet A, Khan R, Zullino D, et al. Internal consistency and measurement equivalence of the cannabis screening questions on the paper-and-pencil face-to-face ASSIST versus the online instrument. Subst Abuse Treat Prev Policy 2015;10:8 [FREE Full text] [doi: 10.1186/s13011-015-0002-9] [Medline: 25886462]

33. Kongsved SM, Basnov M, Holm-Christensen K, Hjollund NH. Response rate and completeness of questionnaires: a randomized study of Internet versus paper-and-pencil versions. J Med Internet Res 2007;9(3):e25 [FREE Full text] [doi: 10.2196/jmir.9.3.e25] [Medline: 17942387]

34. Ritter P, Lorig K, Laurent D, Matthews K. Internet versus mailed questionnaires: a randomized comparison. J Med Internet Res 2004 Sep 15;6(3):e29 [FREE Full text] [doi: 10.2196/jmir.6.3.e29] [Medline: 15471755]

35. Thorén ES, Andersson G, Lunner T. The use of research questionnaires with hearing impaired adults: online vs. paper-and-pencil administration. BMC Ear Nose Throat Disord 2012;12:12 [FREE Full text] [doi: 10.1186/1472-6815-12-12] [Medline: 23107440]

36. 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]

37. Vallejo MA, Mañanes G, Isabel Comeche M A, Díaz MI. Comparison between administration via Internet and paper-and-pencil administration of two clinical instruments: SCL-90-R and GHQ-28. J Behav Ther Exp Psychiatry 2008 Sep;39(3):201-208. [doi: 10.1016/j.jbtep.2007.04.001] [Medline: 17573039]

38. Verkuil B, Brosschot JF. The online version of the Dutch Penn State Worry Questionnaire: factor structure, predictive validity and reliability. J Anxiety Disord 2012 Dec;26(8):844-848. [doi: 10.1016/j.janxdis.2012.08.002] [Medline: 23023163]

39. Vleeschouwer M, Schubart CD, Henquet C, Myin-Germeys I, van Gastel Willemijn A, Hillegers Manon H J, van Os Jim J, Boks Marco P M, et al. Does assessment type matter? A measurement invariance analysis of online and paper and pencil assessment of the Community Assessment of Psychic Experiences (CAPE). PLoS One 2014;9(1):e84011 [FREE Full text] [doi: 10.1371/journal.pone.0084011] [Medline: 24465389]

40. Whitehead L. Methodological issues in Internet-mediated research: a randomized comparison of internet versus mailed questionnaires. J Med Internet Res 2011;13(4):e109 [FREE Full text] [doi: 10.2196/jmir.1593] [Medline: 22155721]

Abbreviations

CDI: Children's Depression Inventory
DIKJ: Depressions-Inventar für Kinder und Jugendliche
ICC: Intraclass Correlation Coefficient
SCAS: Spence Children's Anxiety Scale

XSL•FO

RenderX

Original Paper

Using Visual Analogue Scales in eHealth: Non-Response Effects in a Lifestyle Intervention

Tim Kuhlmann[1], Dipl.-Psych.; Ulf-Dietrich Reips[1], PhD (Psych), MSc (Psych); Julian Wienert[2], PhD (Psych); Sonia Lippke[3], PhD (Psych), Dipl.-Psych.

[1]Research Methods, Assessment & iScience, Department of Psychology, University of Konstanz, Konstanz, Germany

[2]Institute for Social Medicine and Epidemiology, University of Lübeck, Lübeck, Germany

[3]Jacobs Center on Lifelong Learning and Institutional Development, Department of Psychology & Methods, Jacobs University Bremen, Bremen, Germany

Corresponding Author:
Tim Kuhlmann, Dipl.-Psych.
Research Methods, Assessment & iScience
Department of Psychology
University of Konstanz
Universitätsstraße 10
Konstanz, 78454
Germany
Phone: 49 7531882895
Fax: 49 7531882895
Email: tim.kuhlmann@uni-konstanz.de

Abstract

Background: Visual analogue scales (VASs) have been shown to be valid measurement instruments and a better alternative to Likert-type scales in Internet-based research, both empirically and theoretically [1,2]. Upsides include more differentiated responses, better measurement level, and less error. Their feasibility and properties in the context of eHealth, however, have not been examined so far.

Objective: The present study examined VASs in the context of a lifestyle study conducted online, measuring the impact of VASs on distributional properties and non-response.

Method: A sample of 446 participants with a mean age of 52.4 years (standard deviation (SD) = 12.1) took part in the study. The study was carried out as a randomized controlled trial, aimed at supporting participants over 8 weeks with an additional follow-up measurement. In addition to the randomized questionnaire, participants were further randomly assigned to either a Likert-type or VAS response scale version of the measures.

Results: Results showed that SDs were lower for items answered via VASs, $2P$ (Y ≥ 47 | n=55, P=.5) < .001. Means did not differ across versions. Participants in the VAS version showed lower dropout rates than participants in the Likert version, odds ratio = 0.75, 90% CI (0.58-0.98), P=.04. Number of missing values did not differ between questionnaire versions.

Conclusions: The VAS is shown to be a valid instrument in the eHealth context, offering advantages over Likert-type scales. The results of the study provide further support for the use of VASs in Internet-based research, extending the scope to senior samples in the health context.

Trial Registration: Clinicaltrials.gov NCT01909349; https://clinicaltrials.gov/ct2/show/NCT01909349 (Archived by WebCite at http://www.webcitation.org/6h88sLw2Y)

KEYWORDS

XSL•FO
RenderX

Introduction

Internet- and eHealth-Based Health Promotion

Internet-based health promotion targeting health behavior change is a promising approach for health behavior change and aftercare programs [3]. It offers a number of advantages over traditional health promotion methods, most notably, the possibility of taking part anywhere at any time, as long as an Internet-connected device is available. Furthermore, it offers the possibility of tailoring the health promotion to the specific users. Tailored content and feedback can be presented based on demographic characteristics such as gender and age and also based on answers to items presented during the health promotion period or previously assessed behavioral outcomes. Such personalization has been shown to increase the effectiveness of health promotion [3,4]. Initiatives of eHealth may also help in increasing the cost-effectiveness of health care programs, as their running costs are comparably low. The field of Internet-based health promotion is in its early stages and thus, guidelines for designing such programs optimally are sparse. This is of crucial importance because the development of eHealth programs is time consuming and the outcome, for example, less mortality due to a healthier lifestyle, of high value [5].

Dropout has been discussed as a frequent phenomenon in the literature on Internet research (eg, [6,7]). Musch and Reips found in an early meta-analysis that median dropout is at 35% [8]. Reips [9] lists several factors that may influence dropout by participants in Internet-based research: incentives, placement of demographic questions (early means less dropout), technology used (server-side programming reduces dropout), attractive website design, trustworthiness of website, offering feedback, information about duration and progress, and so forth. Eysenbach writes "… factors, for example, expectation management before the trial or 'push factors' such as reminders by the study team, influence the shape and slope (steepness) of the attrition curve." (p.4) [6].

The possibility for participants to drop out at any time during Internet-based studies also translates to tailored eHealth programs. The implications are potentially more serious, however. Dropout in a tailored eHealth and online health promotion program not only means missing data for the researcher but also a less effective treatment for the participant. In general, eHealth aims at lifestyle changes to increase the health and lower the risk factors for participants who oftentimes are already at high risk. Dropping out of the program lowers the chances of behavior change and can therefore have serious consequences for the individuals' health and may sometimes lead to more complications and even an earlier death.

Measures to avoid, reduce, control, or use dropout [9,10] such as the *seriousness check*, *warm-up*, and *high hurdle* techniques to identify participants who just want to take a look out of curiosity (Eysenbach consequentially writes about a "curiosity plateau", p.4 [6]) are even more important in eHealth programs compared with Internet research. This is especially true because dropout is often higher for subgroups with an increased risk for adverse health events. These are participants with the highest

possible gain from adhering to the guidelines and suggestions of the program [11]. Furthermore, *missing data* have severe implications for eHealth as well. Tailoring of the health promotion program is dependent on the data provided by participants. Non-response to certain variables therefore results in fewer tailoring opportunities, thus potentially lowering the effectiveness of the whole program [5].

Visual Analogue Scales

When designing an online questionnaire, the question of how to design items and which answer options to present is always an issue. In the case of measuring the degree of agreement to a statement, the answer options are traditionally often presented in a Likert-type scale format. On these scales, participants have the option to distinguish between a discrete set of answer options to indicate their agreement. They are widely implemented in questionnaires within the medical and social sciences, for example, when measuring subjective health, personality, intentions, or expectancies. Typically, the response to an item is indicated in 5 to 7 gradations ranging from "strongly disagree" to "strongly agree." An alternative to Likert-type scales with radio buttons are *visual analogue scales* (VASs). VASs are rating scales in a continuous graphical format. Instead of clicking on a radio button to indicate the amount of agreement to a statement, participants click anywhere on a continuous line that stretches between the 2 anchors.

These scales offer a number of advantages over Likert-type scales with regard to psychometric properties and data analyses. They were first described by Hayes and Patterson [12]. A theoretical advantage over discrete scales, such as radio button scales, often used on the Internet is that answers are not restricted to a certain number of response options, but rather, very fine gradations can be measured. In computer- and Internet-based assessment, each pixel in length of a VAS corresponds to a possible value [2]. Data collected with VASs offer more options for statistical procedures than data collected with categorical scales, for example, transformations, recoding of scores, and more power for goodness-of-fit-tests. Mean ratings for assessed scales have been shown to be equal when comparing the value received from VASs to the one gathered via other answer scales in paper-based studies [13] and Internet-based studies [1].

In a paper-based study, Myles et al [14] and Myles and Urquhart [15] found that data from VASs form a linear scale when assessing mild-to-moderate pain intensity. Furthermore, equal changes in intensity corresponded to equal differences in length on VASs. Reips and Funke [2] found that in Web questionnaire items, VASs fulfill the requirements of measurement on the level of an interval scale. Therefore, differences between ratings on VASs can be interpreted in a meaningful way, and the prerequisites for many statistical procedures are met. Reips and Funke [2] also examined whether the length of the presented VAS had an influence on the quality of the data provided. They investigated 3 different lengths: 50, 200, and 800 pixels. Despite small differences between the different length versions of the scale, all 3 versions provided very similar data of high quality. Hayes et al [16] described the possibility of general VASs to maximize the discrimination between neighboring categories

and possibly produce less compressed answers at the extremes depending on the anchors used.

Visual Analogue Scales and Missing Data

The impact of VASs on dropout and missing values differs between studies. Couper et al [17] reported a higher dropout in VAS questionnaires (8.2%) in comparison to questionnaires with radio button scales (4.4%). Frequency of missing values was also higher in the VAS questionnaire version: 6.7% as compared with 1% in radio button scales. They note, however, that these differences are confounded by the use of Java applets in the VAS questionnaire version, which resulted in longer loading times and could have caused technical difficulties. Funke and Reips [1] found no difference with regard to item non-response in their study on semantic differentials, comparing VASs and 5-point Likert scales. In fact, they observed a trend toward more complete responding for the VAS questionnaire version.

Tucker-Seeley [18] conducted a Web-based experiment on students in grades 4 to 12, where the students had to answer the same question in a Likert or VAS format. The results showed that 71.4% of participants preferred to answer questions in a VAS format. Furthermore, in the VAS questionnaire version, 76% of respondents indicated that they felt they could pick an answer that matched exactly the way they felt, compared with only 51% in the Likert questionnaire version. Given the influence of motivational factors on dropout [19,20], these findings also suggest a possible positive influence of the format on non-response.

Research Questions and Hypotheses

The main goal of this study was to investigate VASs in the applied context of an eHealth program. We replicated previous research, and expected to find lower standard deviations (SDs) in the VAS questionnaire version [1] but no difference in means between VASs and other scales [1,17]. Furthermore, we hypothesized that the use of VAS would reduce dropout and missing values in comparison to the Likert questionnaire version. Although some research showed the opposite effect [17], more recent studies have found a trend toward less missing data [1]. We argue that the progression in technology has widely eliminated technical difficulties when using VASs, and therefore, a main effect on less dropout and missing data remains, even with (older) cardiac patients who otherwise may not be routinely confronted with computerized VASs.

Methods

Likert vs Visual Analogue Scales

The experimental manipulation of the present study was the use of either Likert-type scales or VAS scales in the questionnaire at the beginning of the study. The 7-point Likert scale ranged from "Strongly Disagree" to "Strongly Agree," and answers were indicated via radio buttons. The VAS used the same labels, but a continuous answer scale was presented instead of radio buttons. Both scales were displayed in approximately the same length on the screen, and the number of items per page was identical in both the questionnaire versions. An example item from the Likert scale and VAS scale formats is shown in Figure 1. The manipulation was applied to all items that measured a degree of agreement to a statement. Items with assumed discrete answer options, for example, stages of change, were excluded from the manipulation.

Characteristics of the eHealth Study

The eHealth program was developed as an aftercare program for cardiac rehabilitation patients. The aim of this program was to provide support for participants in maintaining behaviors learned during cardiac rehabilitation, namely physical activity and a healthy diet. A detailed description of the health promotion program is published in a study protocol [21]. Results investigating the effectiveness of the program were published by Storm et al [22] and Reinwand et al [23]. The program was implemented via a university-external provider of solutions for Internet-based tailored programs. It lasted 8 weeks and took place immediately after the rehabilitation period. The program consisted of an extensive questionnaire at the beginning and end of the period, where behavioral data, social-cognitive variables, and sociodemographic data were assessed. The program was comprised of 8 weekly sessions consisting of different behavior change techniques [21,24]. Depending on indications in the questionnaire and behavioral data during the program, the content was tailored to each participant. Each session was designed to take about 20 minutes to complete. Participants did not have to log in at a specific time but were able to participate when it suited their schedule during the week. Weekly email reminders were sent to every participant (but see 'Dropout across sessions' below.)

Figure 1. Example item from the questionnaire in VAS and Likert formats.

Procedure

'Recruitment for the study took place at 9 cardiac rehabilitation centers in Northern Germany, via local newspaper articles, research agency online panels and information on Jacobs University's homepage. Participants interested in joining the eHealth study were asked to create an account on the external provider's homepage. Creating an account led directly to the baseline questionnaire (*t1*), followed by the first session. Owing to the expected health benefits of the study, as well as providing the control group with an incentive to provide data directly after rehabilitation and after 8 weeks, we decided to make use of a randomized controlled trial (RCT; ClinicalTrials.gov registration number NCT01909349) design with a no-program group (NPG). Participants were randomly assigned to either the program group (PG) or the group that was only asked to fill in the questionnaire (NPG). The study design and the current data, which were analyzed to answer our research questions, are displayed in Figure 2.

In both the groups, participants received questionnaire items either in a 7-point Likert-scale answer format or in a VAS format. This factor was randomly assigned between subjects, with each participant only receiving Likert- or VAS-formatted items. Figure 2 shows the randomization process and flow of participants. The questionnaire at *t1* was the same for the PG and NPG, where the content changed only after the end of the questionnaire. PG participants started with the first session, whereas the NPG received the information that their program would continue 8 weeks later. Six important items for the tailoring were designed as forced choice. However, we designed all other items so they could be skipped without answering because the forced choice format is to be avoided in Internet-based questionnaires for methodological reasons [25,26]. The forced choice items asked for the gender and for a nickname to be chosen for the program. Furthermore, answers were required for 4 items assessing self-efficacy and stages of change.

Figure 2. Randomization of participants to groups and response scale versions. NPG: no-program group, PG: program group.
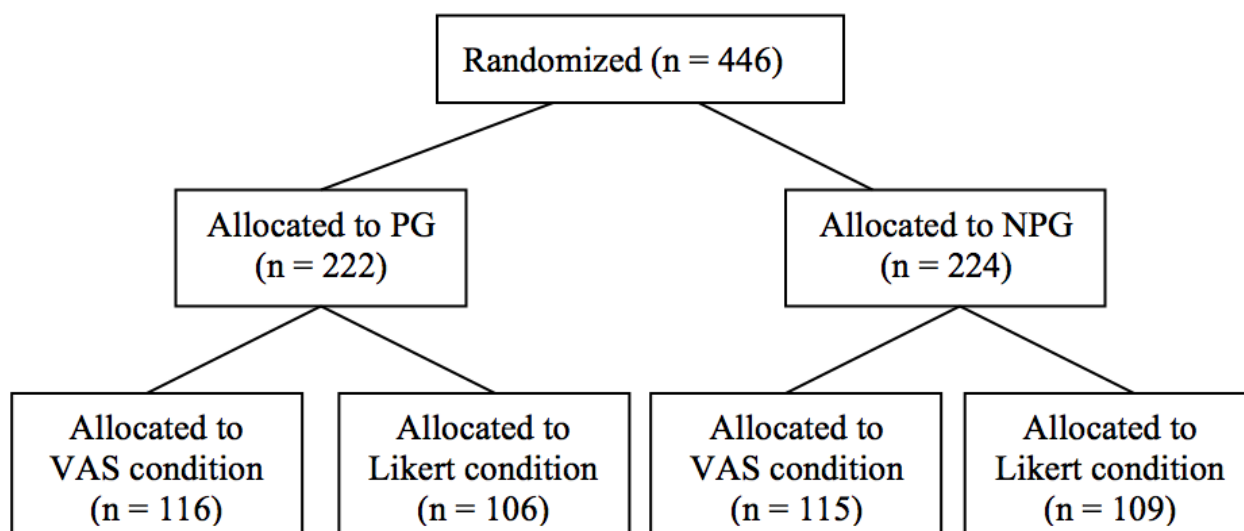
Measures

The questionnaire contained constructs typically assessed in health behavior studies. These included stages of change, risk perception, intentions, outcome expectancies, self-efficacy, action plans, coping plans, social support, and habits [21]. Most of the constructs were assessed separately for physical activity and dietary behavior. Furthermore, health-related variables such as height, weight, physical activity, and dietary behavior were also assessed via self-report. Sociodemographic variables included—among others—gender, education, year of birth,

native language, and marital status. In total, the questionnaire contained 129 items. Of these, 55 items were presented in either a Likert or VAS format.

Dropout was analyzed separately [20,27]. Furthermore, the dropout across the 8 weekly sessions was analyzed. This analysis could only be performed for the PG, as the NPG was not able to participate in the program until the PG finished the 8-week program, diminishing comparability between the groups.

Sample Characteristics

The sample consisted of German cardiac patients who took part in ambulatory or stationary rehabilitation and German participants interested in improving their physical activity and diet. The total sample was comprised of 446 participants, about one third (35%) of whom were from the original target group of medical rehabilitation patients. Almost all participants (97%) were native German speakers. The remaining participants indicated English or Lithuanian as their native language. The ages ranged from 24 to 77 years (M=52.4, SD=12.1). Female participants comprised 59% of the sample. The level of education of participants varied, with 52.6% of the sample having finished high school or having earned a higher degree. Most of the participants (75%) were married or in a relationship.

Results

The results from the analyses of item properties, dropout, and missing values are presented in the following section. In line with Reips and Franek [28], data from participants who dropped out were included in the analyses before the point of their dropout.

Means and Standard Deviations

Answers in the Likert questionnaire version were recoded to make values comparable to the VAS questionnaire version. Recoded values ranged from 1 to 100. T-tests comparing the mean scores for the 55 items that differed between questionnaire versions were then calculated. There were no differences in the means between the questionnaire versions. In line with previous findings on VASs, SDs were smaller for 47 items in the VAS questionnaire version, $2P (Y \geq 47 \mid n=55, P=.5) < .001$. Items in the VAS questionnaire version had a pooled SD of 30, whereas in the Likert questionnaire version, the pooled SD was 33. Latent variables were analyzed using structural equation modeling. The 55 items amounted to 15 different measured

constructs. For these, confirmatory factor analyses were conducted with free estimates for variance and mean of the latent variables. Scale version was used as a grouping variable in the analyses. In none of the 15 cases was there a mean difference between the latent estimates of the different scale versions (all z's<1.96). The variance of the latent variables was smaller in the VAS group in 13 of the 15 studied constructs. These results are similar to previous findings on distributional properties of VAS scores [1,2].

Dropout Across Sessions

The analysis of dropout across sessions was only possible in the PG because the NPG was not invited to weekly sessions until the other group ended their treatment. Between-session dropout was compared across the weekly sessions until the follow-up questionnaire was completed. Overall, there were 9 different sessions with 8 weekly parts, and the follow-up questionnaire was administered 1 week later. Because participants were able to continue participation even if a session was missed, the between-session dropout curves do not monotonically decrease. The dropout within the 2 questionnaire versions is shown in Figure 3. There is a sharp drop in the curve for the Likert questionnaire version from session 1 to session 2. Only 5 of the original 106 participants returned for the second session, indicating a dropout of 95% between the 2 sessions. In the VAS questionnaire version, 64 of 116 participants returned for session 2, which equaled a dropout of 53%. The difference was significant, $\chi^2 (1, N=222) = 63.49, P<.001, \Phi = .54$.

Because such a large difference in dropout due solely to the answer scale used seemed implausible, further possible differences between the questionnaire versions were examined. After careful examination and queries with the external provider of the survey software, it emerged that participants in the Likert questionnaire version did not receive weekly reminder emails, whereas participants in VAS questionnaire version received weekly reminders. This was due to a technical error without options to restore the defect. Aside from the response scale format, this introduces another difference between questionnaire versions across the weekly sessions. As this confounding factor renders the comparison between groups across the weekly sessions inconclusive, dropout was compared across the time frame in which this factor was not present, namely the responses within, and dropouts from the questionnaire during the first session.

Figure 3. Dropout across the weekly sessions until follow-up.
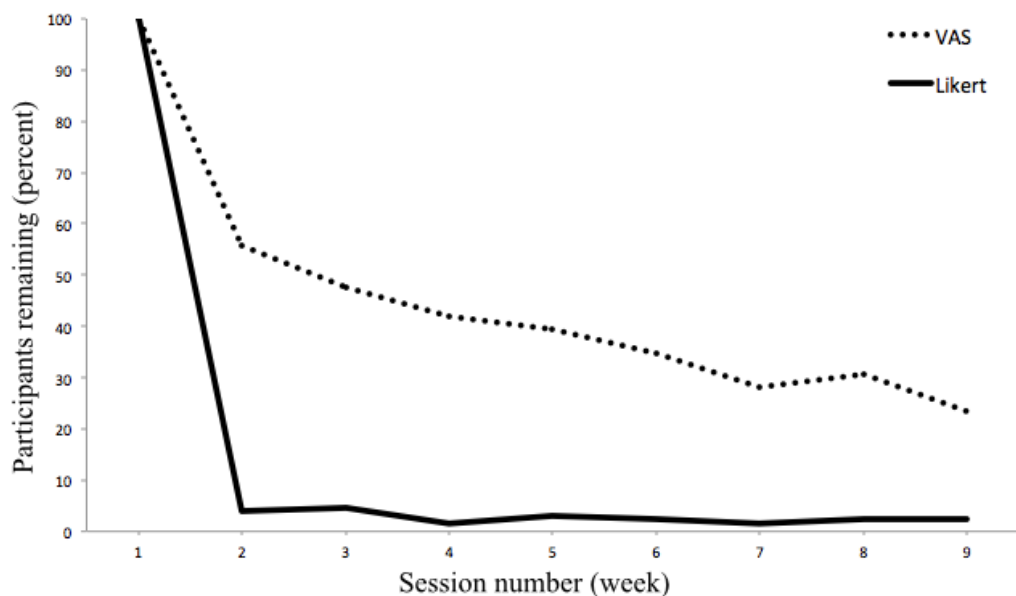
Dropout During the First Session

Dropout was analyzed by determining whether participants were presented a certain item or had already left the questionnaire. Therefore, only skipping a certain page or item was not counted as having dropped out. Figure 4 shows the dropout curves separately for the VAS and Likert questionnaire versions. In the VAS questionnaire version, 73 of 231 initial participants did not finish the questionnaire (31.6%), whereas in the Likert version, 88 of 215 initial participants dropped out (40.9%). A Cox regression with a directed hypothesis resulted in a significant difference between the groups, odds ratio = 0.754, 90% CI (0.58-0.98), P=.04, in favor of the VAS scale to the Likert-scale. Neither gender nor the participation in cardiac rehabilitation before participating in the aftercare program was related to dropout.

Figure 4. Dropout throughout the entire baseline questionnaire by version.
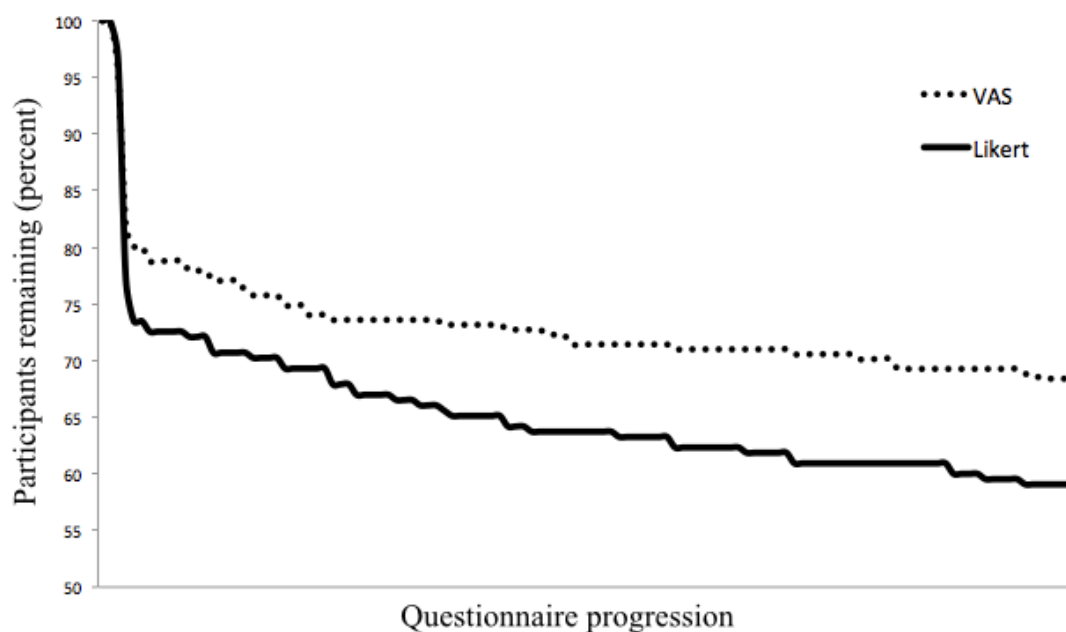
Missing Values

The missing data analysis was performed only for items that were presented as either VAS- or Likert-type answer scales. Items of identical format in both the questionnaire versions were excluded.

The total number of items that could be skipped without answering was 55. We analyzed how many presented items were unanswered by participants. In the VAS questionnaire version, 7% (SD=20%) of the presented items were skipped on average, whereas in the Likert-questionnaire version, 8% (SD=18%) of the items were skipped on average. This difference was not significant. Gender differences in skipped items were also not significant. Age of participants was significantly

correlated with the mean number of skipped items (r=0.26, P<.001), indicating that older participants skipped more questions without answering them.

Discussion

The present study investigated VASs in the applied context of an eHealth program. Lower SDs in the VAS questionnaire version were found for 47 of 55 analyzed items, supporting previous findings that suggest that VAS provide better measurements than Likert scales. There were no differences in mean scores between the VAS and Likert questionnaire versions. Within-session dropout was lower in the VAS questionnaire version. Missing data did not differ between questionnaire versions. Importantly, a technical issue—missing reminder emails in the Likert questionnaire version—was discovered during our investigation and helped identify an important confound, which rendered the between-session dropout analysis inconclusive. The results are discussed in more detail in the following section.

Distributional Properties of VAS

Findings on the mean scores and SDs replicate findings from previous studies in the eHealth context [1,17]. As most results on eHealth studies so far are based on Likert-type scales, mean differences could possibly hinder comparability of results. The replication of previous findings of equal means in the eHealth context therefore provides an important prerequisite for the implementation of VAS response scales in this area. Smaller SDs when using VASs have also been shown by Funke and Reips [1]. A smaller SD translates into smaller standard errors and thus provides more power to statistical analyses [29]. As the recruitment of participants is always difficult in the health context, this is an important reason to consider VASs in this area.

Dropout and Missing Values

In line with our hypotheses and with previous findings in the literature, there was a significant lower dropout in the VAS questionnaire version in comparison to the dropout by participants who were provided with Likert-type scales throughout the first session. A trend toward fewer missing values was also observed. Dropout has been argued to reflect lack of motivation to continue participation, among other factors [6,10,20]. Thus, it is conceivable that the lower dropout when implementing VASs is a manifestation of the preference by participants for this type of scale over Likert-type scales or that study participants became more motivated to remain within the study [18].

The implementation of VASs might result in less reactance and an increased interest in continuing with answering the items. These factors potentially increase motivation and thus result in a lower dropout. The effect sizes for both variables at first seem quite small, a 1% difference in missing values and 9.3% difference in dropout; therefore, further research with larger samples is certainly needed to clarify whether the trend observed in our study is stable in the eHealth context. Given the high importance of retaining participants in programs aimed at reducing the risk of life-threatening cardiac incidents, even small differences in dropout have great implications with regard to health. A 9% difference in dropout means that in the VAS condition, 9% more participants receive a program aimed at improving crucial health behaviors to lower the risk of further cardiac incidents.

VASs have been shown in this study and previous research to possess a number of advantages over other answer formats, for example, better distributional properties and lower standard errors [2,14,15]. With less dropout and no increase in missing values, their use in the eHealth context is even more justified than before.Using VASs in a senior sample with low educational status is a true test of their viability in the eHealth context because it ascertains the validity in this context and for a very specific sample that is represented disproportionately highly in eHealth. Showing a positive effect on dropout and no negative effect on missing values, while also replicating previous findings on advantages, is a strong argument for implementing VASs in the eHealth context. Future research should also investigate the possible effect of VASs on adherence to continuing a study.

The impact of answer scales on dropout across sessions could unfortunately not be determined owing to the confounding factor of missing reminder emails in 1 questionnaire version. Future studies are needed to investigate the impact of VASs on dropout across studies. However, the missing reminder problem would have gone unnoticed without our investigation into dropout differences. This additional finding again highlights the importance of reminders (and checking the reminder functionality), and in the case of cardiac patients, this will eventually mean that the lives of several people may be in a better health condition or even prolonged. Dropout and other types of non-response have long been identified as excellent indicators of technical and motivational issues in Internet-based experiments [19,20]. Our present results provide proof that this diagnostic capability of dropout analysis can and should be extended to Internet-based eHealth studies.

Conclusions

VASs are promising as a viable tool in eHealth studies. Advantages of the format, for example, distributional properties, tend to also hold true in the applied eHealth context. Within-session dropout was lower when VASs were implemented and missing values tended to be lower. Cumulative evidence from several studies points to better measurements with VASs [1,2,14,15]. However, the longer term implications of VASs on between-session dropout will need to be studied, particularly in the context of eHealth.

Care needs to be taken with software design and testing in eHealth programs. It may be better to have the full control that comes with in-house solutions rather than trade in full access for the convenience external providers may suggest.

Acknowledgments

Conflicts of Interest

None declared.

References

1. Funke F, Reips UD. Why semantic differentials in web-based research should be made from visual analogue scales and not from 5-Point scales. Field Methods 2012 May 15;24(3):310-327. [doi: 10.1177/1525822X12444061]

2. Reips UD, Funke F. Interval-level measurement with visual analogue scales in Internet-based research: VAS Generator. Behavior Research Methods 2008 Aug;40(3):699-704. [doi: 10.3758/BRM.40.3.699]

3. Lustria Mia Liza A, Noar SM, Cortese J, Van Stee Stephanie K, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. J Health Commun 2013;18(9):1039-1069. [doi: 10.1080/10810730.2013.768727] [Medline: 23750972]

4. Lippke S, Fleig L, Wiedemann AU, Schwarzer R. A computerized lifestyle application to promote multiple health behaviors at the workplace: Testing its behavioral and psychological effects. J Med Internet Res 2015;17(10):e225 [FREE Full text] [doi: 10.2196/jmir.4486] [Medline: 26429115]

5. Wienert J, Kuhlmann T. A stitch in time saves nine: Things to consider when tailoring your online intervention. European Health Psychologist 2015;17(1):12-17 [FREE Full text]

6. 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]

7. Reips UD. Web-Based Methods. In: Eid M, Diener E, editors. Handbook of multimethod measurement in psychology. Washington, DC: American Psychological Association; 2006:73-85.

8. Musch J, Reips UD. A brief history of web experimenting. In: Birnbaum MH, editor. Psychological experiments on the Internet. San Diego, CA: Academic Press; 2000:61-88.

9. Reips UD. The web experiment method: Advantages, disadvantages, and solutions. In: Birnbaum MH, editor. Psychological experiments on the Internet. San Diego, CA: Academic Press; 2000:89-117.

10. Reips UD. The methodology of Internet-based experiments. In: Joinson A, McKenna K, Postmes T, Reips UD, editors. The Oxford Handbook of Internet Psychology. Oxford, UK: Oxford University Press; 2007:373-390.

11. Moroshko I, Brennan L, O'Brien P. Predictors of dropout in weight loss interventions: A systematic review of the literature. Obes Rev 2011 Nov;12(11):912-934. [doi: 10.1111/j.1467-789X.2011.00915.x] [Medline: 21815990]

12. Hayes M, Patterson D. Experimental development of the graphic rating method. Psychol Bull 1921;18:98-99.

13. Cork RC, Isaac I, Elsharydah A, Saleemi S, Zavisca F, Alexander L. A comparison of the verbal rating scale and the visual analog scale for pain assessment. The Internet Journal of Anesthesiology 2004;8(1):23-38.

14. Myles PS, Troedel S, Boquest M, Reeves M. The pain visual analog scale: Is it linear or nonlinear? Anesth Analg 1999 Dec;89(6):1517-1520. [Medline: 10589640]

15. Myles PS, Urquhart N. The linearity of the visual analogue scale in patients with severe acute pain. Anaesth Intensive Care 2005 Feb;33(1):54-58. [Medline: 15957691]

16. Hayes JE, Allen AL, Bennett SM. Direct comparison of the generalized visual analog scale (gVAS) and general labeled magnitude scale (gLMS). Food Qual Prefer 2013 Apr 1;28(1):36-44 [FREE Full text] [doi: 10.1016/j.foodqual.2012.07.012] [Medline: 23175601]

17. Couper MP, Tourangeau R, Conrad FG, Singer E. Evaluating the effectiveness of visual analog scales: A web experiment. Social Science Computer Review 2006 May 01;24(2):227-245. [doi: 10.1177/0894439305281503]

18. Tucker-Seeley K. The effects of using Likert vs. visual analogue scale response options on the outcome of a web-based survey of 4th through 12th grade students: Data from a randomized experiment [Dissertation]. Boston, MA: Boston College; 2008.

19. Schwarz S, Reips UD. CGI versus JavaScript: A Web experiment on the reversed hindsight bias. In: Reips UD, Bosnjak M, editors. Dimensions of Internet science. Lengerich, DE: Pabst; 2001:75-90.

20. Reips UD. Standards for Internet-based experimenting. Exp Psychol 2002;49(4):243-256. [doi: 10.1026/1618-3169.49.4.243] [Medline: 12455331]

21. Reinwand D, Kuhlmann T, Wienert J, de Vries H, Lippke S. Designing a theory- and evidence-based tailored eHealth rehabilitation aftercare program in Germany and the Netherlands: Study protocol. BMC Public Health 2013;13:1081 [FREE Full text] [doi: 10.1186/1471-2458-13-1081] [Medline: 24245493]

22. Storm V, Dörenkämper J, Reinwand DA, Wienert J, de Vries H, Lippke S. Effectiveness of a web-based computer-tailored multiple-lifestyle intervention for people interested in reducing their cardiovascular risk: A randomized controlled trial. J Med Internet Res 2016;18(4):e78 [FREE Full text] [doi: 10.2196/jmir.5147] [Medline: 27068880]

XSL•FO

RenderX

23.   Reinwand DA, Crutzen R, Storm V, Wienert J, Kuhlmann T, de Vries H, et al. Generating and predicting high quality action plans to facilitate physical activity and fruit and vegetable consumption: results from an experimental arm of a randomised controlled trial. BMC Public Health 2016;16(1):317 [FREE Full text] [doi: 10.1186/s12889-016-2975-3] [Medline: 27066779]

24.   Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]

25.   Stieger S, Reips UD, Voracek M. Forced-response in online surveys: Bias from reactance and an increase in sex-specific dropout. J Am Soc Inf Sci Technol 2007 Sep;58(11):1653-1660. [doi: 10.1002/asi.20651]

26.   Reips UD. Design and formatting in Internet-based research. In: Gosling S, Johnson J, editors. Advanced methods for conducting online behavioral research. Washington, DC: American Psychological Association; 2010:9781433806957-9781433806943.

27.   Bosnjak M. Participation in non-restricted Web surveys: A typology and explanatory model for item non-response. In: Reips UD, Bosnjak M, editors. Dimensions of Internet science. Lengerich, DE: Pabst; 2001:193-208.

28.   Reips UD, Franek L. Mitarbeiterbefragungen per Internet oder Papier? Der Einfluss von Anonymität, Freiwilligkeit und Alter auf das Antwortverhalten [Employee surveys via Internet or paper? The effect of anonymity, voluntariness and age on response behavior]. Wirtschaftspsychologie 2004;6(1):67-83. [doi: 10.5167/uzh-19693]

29.   Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [Medline: 19565683]

Abbreviations

NPG: no-program group
PG: program group
RCT: randomized controlled trial
SD: standard deviation
VAS: visual analogue scale

XSL•FO
RenderX

<u>Original Paper</u>

Evaluation Methods for Assessing Users' Psychological Experiences of Web-Based Psychosocial Interventions: A Systematic Review

Jacqueline Susan Feather[1], PhD; Moira Howson[2], B Hlth Sc Hons; Linda Ritchie[2], B Hlth Sc; Philip D Carter[3], PhD; David Tudor Parry[4], PhD; Jane Koziol-McLain[5], RN, PhD

[1]Centre for Interdisciplinary Trauma Research, Department of Psychology, Auckland University of Technology, Auckland, New Zealand

[2]Auckland University of Technology, Auckland, New Zealand

[3]Auckland University of Technology, School of Computer and Mathematical Sciences, Auckland, New Zealand

[4]Auckland University of Technology, Department of Computer Science, Auckland, New Zealand

[5]Centre for Interdisciplinary Trauma Research, School of Clinical Sciences, Auckland University of Technology, Auckland, New Zealand

Corresponding Author:
Jacqueline Susan Feather, PhD
Centre for Interdisciplinary Trauma Research
Department of Psychology
Auckland University of Technology
A-12
Private Bag 92006
Auckland, 1142
New Zealand
Phone: 64 99219999 ext 7693
Fax: 64 99219780
Email: jackie.feather@aut.ac.nz

Abstract

Background: The use of Web-based interventions to deliver mental health and behavior change programs is increasingly popular. They are cost-effective, accessible, and generally effective. Often these interventions concern psychologically sensitive and challenging issues, such as depression or anxiety. The process by which a person receives and experiences therapy is important to understanding therapeutic process and outcomes. While the experience of the patient or client in traditional face-to-face therapy has been evaluated in a number of ways, there appeared to be a gap in the evaluation of patient experiences of therapeutic interventions delivered online. Evaluation of Web-based artifacts has focused either on evaluation of experience from a computer Web-design perspective through usability testing or on evaluation of treatment effectiveness. Neither of these methods focuses on the psychological experience of the person while engaged in the therapeutic process.

Objective: This study aimed to investigate what methods, if any, have been used to evaluate the in situ psychological experience of users of Web-based self-help psychosocial interventions.

Methods: A systematic literature review was undertaken of interdisciplinary databases with a focus on health and computer sciences. Studies that met a predetermined search protocol were included.

Results: Among 21 studies identified that examined psychological experience of the user, only 1 study collected user experience in situ. The most common method of understanding users' experience was through semistructured interviews conducted posttreatment or questionnaires administered at the end of an intervention session. The questionnaires were usually based on standardized tools used to assess user experience with traditional face-to-face treatment.

Conclusions: There is a lack of methods specified in the literature to evaluate the interface between Web-based mental health or behavior change artifacts and users. Main limitations in the research were the nascency of the topic and cross-disciplinary nature of the field. There is a need to develop and deliver methods of understanding users' psychological experiences while using an intervention.

Introduction

Internet-Delivered Health Care

The past 15 years have included a burgeoning in the development of Internet-delivered health care. The term eHealth has been widely adopted to describe the application of Internet and communication technology to improve the operation of the health care system and health care delivery [1]. The rationale for offering health services online is similar to the provision of any service online, be it commerce, education, or entertainment. It is accessible anytime and anywhere to an almost unlimited audience and is often cheaper than face-to-face [1]. Within health care, the growth of technology means that eHealth is seen as a means of providing targeted treatment to a wide client base [2,3], and self-help Web-based interventions can also contribute to the reduction of disparities in health care internationally [4,5]. It also allows for a focus on patient-centered customized care "providing the right information, to the right person, at the right time" [6]. Moreover, people now expect to be able to use the Internet to assist in most aspects of their lives, including health care.

The range of health services offered over the Internet encompasses not only physical concerns, such as high blood pressure or diabetes, but also psychological issues such as depression, addiction, and anxiety [7-9]. The provision of eHealth for psychological issues can take many different forms. They include (1) direct human interaction, known as e-counseling or e-therapy, which involve electronic communication (via email, Skype, or messaging for example) between a clinician and client; (2) Web-based psychological interventions such as w the new Zealand government's depression support site[10] or Sleepio [11], which have information, health assessment screening, and treatment and are targeted to specific issues; (3) Internet-based therapeutic software programs such as gaming and virtual worlds; (4) mobile apps; and (5) other activities such as support discussion groups, blogs, and podcasts [3,12,13].

It makes sense to assume that the proliferation of Web-based psychological interventions could not have occurred without strong supporting evidence, yet there is a notable gap between what is proposed by eHealth interventions and what is actually delivered [14]. As such, the continued investment in, and adoption of, Web-based interventions requires more evaluative information to ensure resources are not wasted on ineffective interventions and that the "best practices of successful programs are rapidly disseminated" [15]. Evaluative information should be relevant to policy makers, health care providers, practitioners, and clients [16,17]. This requires a multifaceted approach, coalescing aspects of traditional psychological intervention evaluations such as efficacy, effectiveness, and measures of the therapeutic relationship, client variables, and therapist feedback, with computer science developed usability testing approaches to gather an overall measure of an intervention's utility.

Evaluating Web-Based Psychosocial Interventions

It makes sense that evaluation of a Web-based eHealth intervention should naturally follow much of the same evaluative process as a traditional intervention [18]. Effectiveness can be measured by pre- and postintervention assessments using clinical trials. However, assessments can also be made of the components of the intervention. In traditional face-to-face treatment, this would include the therapeutic relationship and session outcome measures, which gather information on the experience of the user during as well as after the intervention. In eHealth, user experience is often focused on usability testing that generally measures the degree to which the intervention is understandable and easy to use.

Effectiveness of Web-Based Psychosocial Interventions

Effectiveness refers to how well a health or psychological treatment works in a real-world setting [19]. Methods used for clinical trials of pharmacological treatments can be applied to complex nonpharmacological interventions such as eHealth initiatives [20]. Randomized clinical trials (RCTs) have found that Web-based interventions are effective in treating a range of psychological issues. For example, individually tailored Internet-delivered cognitive behavioral therapy (CBT) has been found to be an effective and cost-effective means of treating patients with anxiety disorders [9]. A recent meta-analysis of RCTs investigating the effectiveness or efficacy of Web-based psychological interventions for depression found positive results across diverse settings and populations [21]. However, as Richards and Richardson [21] note, effectiveness is only one part, albeit an important one, of an intervention that should be evaluated. For example, effectiveness research does not necessarily investigate the influence of therapist factors.

User Experience of Web-Based Psychosocial Interventions

In evaluation research of face-to-face psychosocial interventions, findings have shown that an effective therapeutic relationship in and of itself may be enough to provide successful outcomes for clients [22]. Moreover, this correlation between ratings of the therapeutic relationship and outcome seems unaffected by other variables such as outcome measure or type of intervention employed [23]. These variables are client derived, based on their subjective experiences, and can be adjusted for and evaluated throughout the treatment program or therapy session. As a result, the interaction between the client and clinician is fine-tuned and dynamic, and the clinician is able to constantly monitor the therapeutic relationship and make changes accordingly. This is possible in a dynamic person-to-person interaction but it is not possible in a one-way interaction between a person and a computer.

In the person-computer interaction, the psychological experience of the person using the therapeutic intervention is lost and cannot be responded or adjusted to (although there are developments in this area [24]). This raises three key concerns. First is the

potential loss of the therapeutic relationship. Second, the psychological experience of the person may be important to determining usage and adherence to an intervention with some reports of dropout rates exceeding 98% from open-access websites dealing with psychosocial issues [25]. Third, it may be important in understanding issues of psychological safety and well-being for "if information is too complex to understand, especially under periods of duress or high cognitive load" then the intervention may not be as effective [6]. Therefore, the in situ psychological experience of the user may be an important variable to be considered; however, it is not clear how it has been or should be evaluated.

Methods of Evaluating Web-Based Psychosocial Interventions

User-Focused Evaluations

In traditional face-to-face interventions, evaluation of user experience is relatively straightforward. For starters, the client can provide instant and direct feedback to the therapist during a session. In addition, there are a number of psychometrically reliable and valid standardized tools such as the Helping Alliance Questionnaire [26] or the Outcome Rating Scale and Session Rating Scale [27], which directly address user experience. In e-therapy or e-counseling, when there is still a person-to-person relationship via Skype or email, the measurement of the therapeutic process is similar as only the delivery differs [12]. However, when the intervention is provided by a website and is driven by computer-programmed algorithms and automated responses it is much more complex. A cross-disciplinary approach is needed that links the psychosocial approach with user-centered design from computer science [6]. One of the most common ways of evaluating the relationship the user has with a website or Web-based intervention is through usability testing.

Intervention-Focused Evaluations

The literature on human-computer interaction yields a number of different methods of evaluating a website, be it a psychosocial intervention or an web-based retail store. One of the most widely used and recognized means of addressing this is usability evaluation [28]. Usability testing is a means of evaluating the design and functionality elements of a website as they apply to the user. However, the end user is not always involved and, if so, any information capturing the user experience is designed to enhance the performance of the intervention in terms of content, design, and navigability of the website, rather than an investigation of the relationship between the user and computer.

Usability inquiry on the other hand is designed to gather broader and more subjective information from the user with regard to the website or intervention. Zhang [29] defined usability inquiry quite broadly as gaining "information about users' experience with the system by talking to them, observing them using the system in real work (not for the purpose of usability testing), or letting them answer questions (verbally or in written form)" [29]. Therefore, although usability testing provides output from the person using the system, the focus is on functionality of the website. In usability inquiry, there is a subtle shift to understanding the users' goals, context, profile, feelings, and thoughts during the process of interaction. The output extends beyond issues of website design and functionality to broader concerns of purpose and context [28].

The methods of ethnographic interviewing, contextual inquiry, cognitive interviewing, and situated co-inquiry lend themselves, somewhat naturally, to an exploration of how users may experience a Web-based psychological intervention. They all place the user as the most important piece of the process. Thus, it fits the paradigm of interpretative and participatory research with elements of phenomenological inquiry to gain an understanding of the users' experience in a natural setting. In some ways, this would come close to understanding the nuance of the relationship between a user and an intervention in a similar way that an evaluation of the therapeutic relationship captures a client's perception and experience of traditional therapy. However, it is unclear to what extent usability inquiry, or any other methods, have been used to address the psychological experiences of those using Web-based psychosocial interventions. Thus, a systematic literature review was undertaken to assess what methods, if any, have been applied to understanding users' in situ *psychological* experience with Web-based interventions.

Methods

Systematic Literature Review Methodology

A systematic literature review was employed to determine how end users' psychological experiences of Web-based interventions have been evaluated. A systematic review is beneficial for this relatively new topic for "identifying gaps and weaknesses in the evidence base and increasing access to credible knowledge" [30,31]. The AMSTAR (Assessment of Multiple Systematic Reviews) was selected [31] to guide the review. AMSTAR factors and how they were applied in this study are outlined below. The research question under investigation was defined as "What evaluation methods have been used to explore end-users' in-situ psychological experiences with web-based psychosocial interventions?"

Inclusion and Exclusion Criteria

There were 4 criteria components for selecting studies. First, the *health issues* of interest were those with a psychosocial component for which a person may have been reasonably expected to see a clinician if the Web-based intervention was not available; for example, helping people with common mental health concerns such as stress, anxiety, and depression. Also included were psychosocial or behavior change intervention programs associated with physical ailments, such as rehabilitation and psychological well-being following surgery or living with human immunodeficiency virus infection. The psychosocial element was required to underscore the importance of the user's psychological experience with the intervention. Thus, interventions with a purely physical component such as smoking cessation were excluded.

Second, the *type of interventions* included were Web-based interventions with a primarily self-help basis that required users to work their way through a series of steps. Excluded were e-counseling and e-therapy programs where the treatment was

based on person-to-person interaction rather than a person-to-computer interaction. They were also excluded if they were solely psychoeducation or health promotion websites that did not require an interactive, self-guided program. Online tools (such as "Facebook") that may be used as part of social skills therapy but not designed for the purpose of treatment or psychosocial intervention were excluded.

Third, the *method* of evaluation had to involve end users' assessment of their experience, rather than expert assessments. Finally, the *focus* of the evaluation must be the end users' psychological experience interacting with the intervention. Studies that looked at the overall efficacy or treatment outcome of a Web-based intervention were excluded. Usability studies were excluded if they focused solely on functional Web-development issues (content, layout, design), rather than on the psychological experience of the user. For example, "think aloud" (also called talk aloud or cognitive interviewing) studies that did not elicit psychological insight or experience from the user but discussed use of menu options, friendliness, and ease of use were not included. Basically, the study had to be user focused, rather than intervention or computer focused. Expert usability and heuristic studies were also not of value as they did not address the actual experience of the intended user of the intervention. If an evaluation of the users' psychological experience formed part of a broader study on the effectiveness study or feasibility study, it was included. As a result of a scoping exercise, we expanded the search protocol to include end users' psychological experience per se, whether in situ or postintervention.

Search Protocol

Keywords applied in the search were determined by an initial review of the literature and modified by feedback from psychologists, computer science usability experts, and a librarian. Sets of terms were proximally searched including experience ("experience," "evaluation," "usability"), Internet ("internet," "web*," "online," "eHealth"), health ("mental health," "behavio*," "psycho*"), and interventions ("intervention," "treatment," "therapy"). To ensure coverage of the health, psychological, and computing components of the topic the following databases were searched: CINAHL (EBSCO), MEDLINE (EBSCO), and Psychology and Behavioral Sciences Collection (EBSCO), Computers and Applied Sciences Complete (EBSCO), ABI/INFORM Global, and IEEE Xplore (see Multimedia Appendix 1). The search was restricted to studies from 2004 to October 2015 because of the relative newness of the field of Web-based interventions. A title scan across non–peer-reviewed journals elicited no studies of interest in the initial scoping, so only scholarly journals were included in the search. A total of 2 studies came from a review of the references of a meta-synthesis of user experience of online computerized therapy for depression and anxiety [32].

Search Results

The search protocol resulted in 62 records that were screened by 2 reviewers (LM and LR), with disagreements resolved by the research team (see Figure 1). After final full-text review, 21 studies were identified as meeting the 4 inclusion criteria. Included and excluded studies are listed in Multimedia Appendices 2 and 3, respectively.

Quality Assessment and Data Synthesis

An assessment of the type and quality of the methods used in the 21 studies was integral to this research project, and screening studies out on the basis of methodological quality would have been counterproductive. Therefore, all the 21 selected studies were included in the synthesis phase. As the identified studies were heterogeneous, a narrative synthesis was undertaken. The a priori data extraction framework involved first, describing the studies (see Multimedia Appendix 2 for study characteristics) and second, their methods of examining user psychological experience of Web-based psychosocial interventions (see Table 1).

Figure 1. Flow Diagram of Study Selection.
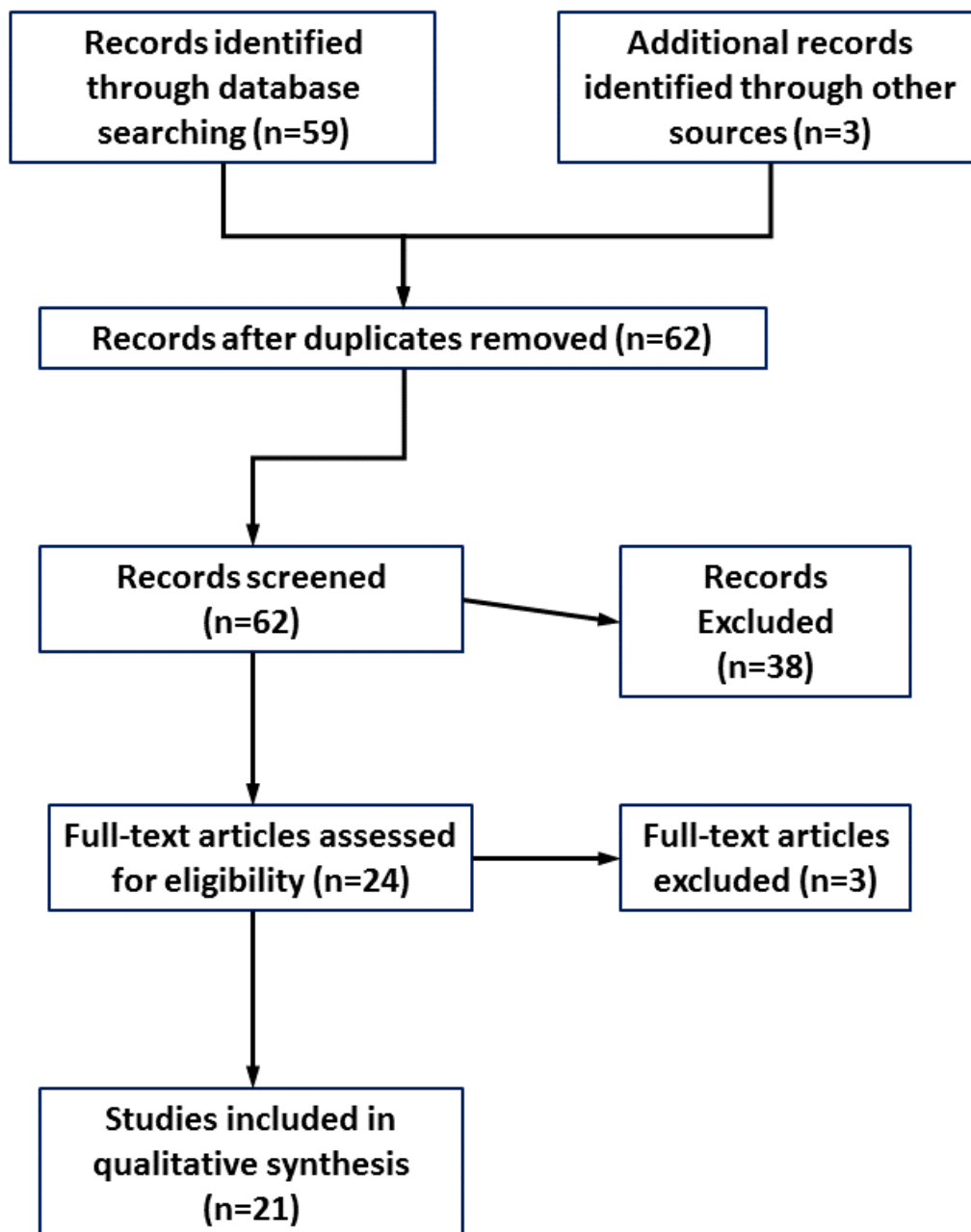
Table 1. Included studies: methods of examining user psychological experience.

Reference	User experience focus	Methods[a]	Tools	Time of assessment
Baños et al [33]	No	Open-ended questions (online)	Researcher designed questionnaire	Postsession (6 × weekly sessions)
Bendelin et al [34]	Yes	Semistructured interviews (face-to-face)	On the basis of the Client Change Interview	Posttreatment (8-10 months)
Bradley et al [35]	Yes	Semistructured interviews (phone)	Researcher devised based on Theory of Planned Behavior	Posttreatment (1 week)
Cartreine et al [36]	No	Questionnaires (online)	System Usability Scale; Credibility Questionnaire; Assessment of Self-Guided Treatment	Postsession
de Graaf et al [37]	No	Questionnaires (online)	Credibility Expectancy Questionnaire (CEQ); customized questionnaire	CEQ at baseline; questionnaire posttreatment (3 months)
Devi et al [38]	Yes	Semistructured interviews (face-to-face)	Researcher devised interview guide	Posttreatment (6 weeks)
Fergus et al [39]	No	Questionnaire; semistructured interviews (face-to-face)	Treatment Satisfaction Questionnaire; researcher devised interview guide	Posttreatment (time unspecified)
Gega et al [40]	Yes	Questionnaire; semistructured interviews (face-to-face)	Session Evaluation Questionnaire; Session Impacts Scale; Helpful Aspects of Therapy form; Client Change Interview	Questionnaires postsession; Interview posttreatment (time unspecified)
Gerhards et al [41]	Yes	Semistructured interviews (face-to-face)	Researcher devised interview guide	Posttreatment (time unspecified)
Gorlick et al [42]	Yes	Semistructured interviews (phone)	Researcher devised interview guide	Posttreatment (<2 years)
Gulec et al [43]	No	Questionnaires (online)	Researcher designed self-report questionnaires	Online weekly during treatment and posttreatment
Hind et al [44]	Yes	Session evaluation forms; semistructured interviews (phone and face-to-face)	Researcher devised session evaluation forms and interview guide	Evaluation postsession; brief interview postsession 1; interview posttreatment (after completing or withdrawing)
Lara et al [45]	No	Questionnaire (online)	Researcher devised questionnaire	Posttreatment
Lederman et al [46]	No	Semistructured interviews (face-to-face)	Researcher devised interview guide	Posttreatment (time unspecified)
Lillevoll et al [47]	Yes	Semistructured interviews (face-to-face)	Researcher devised interview guide based on phenomenological hermeneutical approach	Posttreatment (time unspecified)
McClay et al [48]	Yes	Semistructured interviews (phone)	Researcher devised interview guide based on motivation, experience, and comparison with other treatments	Posttreatment (time unspecified)
Serowik et al [49]	No	Think aloud usability; questionnaire	Think aloud usability protocol; researcher designed questionnaire; modified Working Alliance Inventory	Usability during session; questionnaire posttreatment or at dropout (time unspecified)
Tonkin-Crine et al [50]	Yes	Unstructured interviews (phone)	Open-ended interview researcher devised	Posttreatment (time unspecified)
Topolovec-Vranic et al [51]	No	Unstructured interview (phone)	Unspecified	Weekly during 6-week program and 12 months postenrollment
Van Voorhees et al [52]	No	Questionnaire; diaries	Researcher designed questionnaire	Diary during and after session; questionnaire posttreatment (time unspecified)
Wade et al [53]	No	Semistructured interviews; survey	Unspecified	At follow-up (time unspecified)

[a] Other measures may have been used in the study such as pre- and postbaseline measures of diagnosis but these were not included in the data extraction as they did not concern user experience.

XSL•FO
RenderX

Results

Overview of Included Studies

Intervention Type

The interventions were Web-based self-help interventions consisting of a number of modules (ranging from 3 to 12), most expected to be used on a weekly basis. Although all of the studies were predominantly self-help programs, some were augmented by other elements such as an online diary [37,38,41], social or peer group support in the form of discussion or chat groups [38,39,42,44,46], and some degree of therapist or tailored guidance to support the self-help component [34,38-40,43,53]. A total of 9 studies focused on an intervention in the pilot or development phase [33,34,36,39,43,46,49,52,53], whereas 13 studies focused on existing treatment programs such as "MoodGYM" or "Colour Your Life" [35,37,38,40-42,44,45,47,48,50-52]. All of the interventions contained interactive or homework components and some reported including multimedia such as audio or video in content delivery.

Of the 21 studies, most studies focused on the treatment of psychological issues. A total of 10 studies evaluated interventions designed to treat or prevent the development of depression using principles of CBT as the modality of treatment [34,36,37,40,41,44,45,47,51,52]. One intervention used CBT to assist adolescents with stress, anxiety, and depression [35]. Of the studies, 2 focused on delivering mainstream treatment for eating disorders online [43,48] and 1 on the treatment of first psychotic episodes through the use of CBT, support, and psychoeducation [46]. One intervention used motivational interview principles to support veterans with war exacerbated psychiatric issues apply for jobs and benefits [48]. Another aimed to reduce child behavior problems and parenting stress for young children with traumatic brain injury (TBI) via an online self-guided program and live coaching [53]. A total of 4 studies focused on interventions delivering psychological well-being and support to those with chronic or acute physical issues including irritable bowel syndrome [50], cancer [39,42], and angina [38]. Another study used a positive psychology approach to enhance mood in general populations [33].

Study Characteristics

Of the 21 studies, 13 were published from 2012 onward (see Table 1). A total of 10 studies were explicitly focused on the evaluation of the patient or users' psychological experience of the intervention [34,35,38,40-42,44,47,48,50]. Of the remaining 10 studies, 8 gathered user experience as part of the development and pilot testing of the intervention [36,39,43,45,46,49,52,53]. Another study collected information on user experience to assess the acceptability and use of online CBT for depression [37], and another looked at the efficacy of using an online CBT intervention for depression with patients experiencing depression along with their TBI [51].

Evaluation of User Experience

Rationale for Evaluating User Experience

The rationale for evaluating user experience varied across the studies. Of these, 5 studies were seeking information on the users' psychological experience of using the website to inform improvements [36,39,45,46,49]. A total of 6 studies were interested in finding out how to increase use of online interventions by understanding barriers to change [35] and acceptability in particular populations [42-44,51,52].

Process of Evaluating User Experience

The process of gathering user experience information also varied. A total of 13 studies relied on only one form of evaluation method. Of these, 10 studies used only an interview at the end of the treatment program (with a range of 1 week to 2 years posttreatment, where specified) to ask users about their experience with the intervention. Of the studies, 5 used face-to-face interviews [34,38,41,46,47] and 5 employed phone interviews [35,42,48,50,51]. The interviews were conducted by the researcher or research assistants (when specified) and their duration ranged from 19 minutes [35] to 111 minutes [34]. Of the studies, 3 employed only questionnaires to investigate user experience. Of the 3 studies, 2 deployed it at the end of each treatment session [33,36] and the other [37] at the beginning of treatment and then 3 months posttreatment.

The remaining 8 studies relied on a mixed methods design to investigate user experience. Of these, 4 studies employed a posttreatment semistructured interview, 1 study [39] complimenting it with a posttreatment questionnaire and another study with a posttreatment satisfaction survey [53]; Gega and colleagues [40] with postsession questionnaires and Hind and colleagues [44] with a postsession evaluation form. One study used weekly evaluation questions and a postintervention questionnaire [43]. A further study used a during and after session diary to capture user feedback combined with a posttreatment questionnaire [52] and another used module and a postintervention evaluations along with content analysis of the site's discussion forums [45]. The final study used a think aloud usability process during the session followed by an end-of-treatment questionnaire [49]. Note that this was the only study employing a usability testing method that included a user experience focus, explicitly asking for feelings during the session, and a questionnaire to elicit user experience with the therapeutic alliance during the Web-based intervention.

Tools Used to Evaluate User Experience

Tools used to gather user experience information ranged from customized to off-the-shelf. Of the 14 studies that used some form of interview with users, 12 were semistructured with the topic guide designed by the research team. Of these, 2 studies provided a theoretical basis for the design of the interview guide. Lillevoll and colleagues [47] followed a phenomenological-hermeneutic approach to understand the lived experience of the user of the intervention in the natural setting, and Bradley and colleagues [35] followed a Theory of Planned Behavior. A further 2 studies [34,40] based the interview guide on the Client Change Interview (CCI), which

was designed as an empathic exploration of aspects of a client's experience with traditional face-to-face counseling and therapy.

Questionnaires used to evaluate user experience were either researcher designed or existing tools. Of the studies, 7 used self-customized questionnaire or evaluation forms [33,37,43-45,52,53]. Others relied on questionnaires designed to gain feedback about the therapeutic process, sessions, expectations, and outcomes (Table 2).

Table 2. Treatment feedback questionnaires used to evaluate user experience.

Questionnaire	Purpose	Cited in[a]
Credibility Questionnaire	Credibility of computer programs, psychotherapy, and treatment	Cartreine et al [36]
Assessment of Self-guided Therapy	Acceptability of treatment (eg, comfort, personal acceptance)	Cartreine et al [36]
Credibility Expectancy Questionnaire	Expectation and rationale for treatment	de Graaf et al [37]
Treatment Satisfaction Questionnaire	Satisfaction and experience (convenience, quality, value) with online treatment	Fergus et al [39]
Session Evaluation Questionnaire	User experience with the session in terms of depth, positivity, smoothness, and arousal	Gega et al [40]
Session Impact Scale	User view of session impact on understanding, problem solving, therapeutic relationship, and hindering impact	Gega et al [40]
Helpful Aspects of Therapy	Identify helpful or hindering aspects of the treatment session	Gega et al [40]
Working Alliance Inventory	Self-report assessment of user experience of alliance with treatment (modified for online)	Serowik et al [49]

[a] Further details of questionnaires can be found by consulting the references.

Analysis of User Experience Data

As this study is focused on the methods of evaluating user experience as opposed to user experience per se, the results of the studies are not presented in detail. In summary, quantitative analysis was carried out on the statistical data gathered in the questionnaires and descriptive summaries reported where appropriate. The qualitative user experience information in 19 of the studies was analyzed using content analysis to elicit emerging themes. Of these, 2 studies followed elements of grounded theory methodology to develop theories of user experience with Web-based interventions [41,52]. One study took a phenomenological-hermeneutic approach [47].

Overall, these studies reported thematic categories to describe the users' experiences within the frame of the study purpose. Thus, the themes that evolved were varied and included user experience in terms of the therapeutic process [40,41], individual and social aspects of the intervention [33,40-42,47], as well as Web-based characteristics of the intervention [40-42,48-50]. Process issues included motivation and user profile [34], privacy and help-seeking [35], overall experience [37,46], barriers [50,52], expectations [37], personalization, solitude, and social, individual, and contextual features [40,41]. For example, in one of the grounded theory studies, Gerhards and colleagues [41] identified themes of the user experience in terms of the computer, individual, social, research, and environmental contexts.

Discussion

Key Findings

This systematic review identified a gap in examining users' *psychological* experience of Web-based interventions in situ. Among studies using usability testing, only 1 study [49] explicitly explored what it felt like for the user in-session. Most were focused on the functionality of the intervention in terms of Web design, ease of use, and readability. Usability inquiry methods such as situated co-inquiry proposed by Carter [54] or ethnographic and contextual inquiry discussed by Cooper et al [28] are not yet evident in the literature. It is important that our methods of evaluating Web-based interventions include opportunities for uncovering the relationship between the user and the computer as a social agent. The studies that did focus on the therapeutic experience online tended to evaluate it postsession or posttreatment using tools designed for evaluating face-to-face interventions. For Web-based interventions, as called for by Knowles et al, "it is likely that more in-depth, observational data collection methods will be necessary to better capture user experience in [the] future" (p11) [32].

The rationale for postsession and posttreatment evaluation occurred when the interest was social science rather than computer science and the focus was the treatment experience rather than the mode in which it is delivered. Some of the studies sought to understand the therapeutic process at a session level [40,44], the therapeutic relationship [34,40], and how overall experience translated from traditional delivery channels of treatment into Web-based treatments [41,43]. Other studies aimed to enhance the treatment program and increase usage and acceptability of particular populations to certain interventions, such as adolescents with psychological distress [35] or those with psychosocial issues associated with physical health problems [38,39,42,44,50,51]. Postsession and posttreatment evaluations were also used to understand barriers to usage [52] and motivations to participate in Web-based interventions [37,48]. Most of these evaluations used semistructured interviews or questionnaires to collect information. Posttreatment semistructured interviews (face-to-face or phone based) were the key methods used to understand the experience of the user.

An in-depth review of the methods was not possible because of gaps in the literature on timing, length, interviewer characteristics, and topic guide details. It is evident however that the problem with posttreatment evaluations is the potential time delay between treatment and recollection of experience and this potentially affects validity and reliability of the information. The time between treatment and evaluation varied, some took place 1 week posttreatment [35] and another had a gap of up to 2 years [42]. Some studies did not actually specify how long following treatment that the evaluations occurred. Some interviews were conducted by phone for 20 minutes [48] and others reported almost 2-hour face-to-face interviews [34] and thus the depth and quality of information varied accordingly. In addition, the characteristics of the interviewers were also sometimes overlooked. When specified, the interviews were conducted by the researcher or research assistant who may or may not have had relevant clinical training. In general, the topic guides or questions asked by the researchers were not published so that the areas discussed with the participants are unknown, although it can be extrapolated to some extent by the focus of the research. Some studies provided example topics or questions. For example, Devi and colleagues [38] presented a table of example interview questions that included "What were your initial thoughts and feelings regarding this web-based programme?" and "What was your overall experience of using the programme?" Having more information as to the topics covered would be useful to understand in detail what aspects of user experience were investigated.

In addition to interviews, questionnaires were also used; most commonly they were standardized tools to measure experience with traditional treatment or therapy modified for eHealth. For example, Gega and colleagues [40] used the Session Impact Scale, Session Evaluation Questionnaire, and Helpful Aspects of Therapy questionnaire to gather immediate feedback from users following each online session. This elicited in-depth session and overall feedback that were categorized into a number of themes, such as the experience of "learning by doing" or having no fear of being judged by a Web-based intervention. These questionnaires were designed to understand treatment experience and outcomes and were not concerned with the functionality of the website. The use of the CCI as the basis for the interview guide by Bendelin and colleagues [34] and Gega and colleagues [40] reinforced this focus on treatment versus delivery that was also evident in the analysis for the evaluation findings. The analysis of user feedback employed by these studies reflected an interest in understanding users' psychological experience for its own sake rather than in relation to the intervention alone. The value of the information elicited was dependent on the purpose of the research, and each of the 21 selected studies had slightly different foci and rationale for exploring the experience of the user.

Most of the studies that used an interview approach used inductive content analysis to interpret the findings in terms of themes and subthemes of experience. Examples of themes included user profile and motivation [34], privacy, and help-seeking [35]; solitariness and personalization [40]; and the individual, social, computer, and wider contexts of experience [41]. Interviews are widely used in information systems research

and eHealth evaluation, but it is important to be clear that an interview is an artificial and additional intervention. Myers and Newman's [55] description of the dramaturgical model as originally described by Goffman [56] may aid interviewers to move between evaluation and clinical issues.

Although functionality was not a driver in the research, themes of intervention characteristics [36,40,41] and Web engagement [47] did emerge. This suggests that users' psychological experiences are determined in part by the psychological issue and treatment process, and in part by the channel or mode of delivery. In other words, the functionality explored in usability testing might be a piece of the experience but is by no means all of it.

The choice of methods of evaluation depends primarily on which aspect of the intervention is most critical to measure—for example, effectiveness, usability or engagement—and on the maturity of the implementation of the intervention, and multiple methods may be appropriate. For early prototypes, interviews directly after the intervention are appropriate in order to gain an understanding of all of the issues and potential benefits as seen by the users. Larger-scale interventions will be more likely assessed by questionnaires, often modified from existing assessments. However, these should occur as quickly as possible after the intervention so that the users still have the experience fresh in their minds. Standard functional questionnaires may be preferred when Web-based treatment approaches are being compared with other interventions. However, all of these evaluation approaches must bear in mind the "emergent" nature of eHealth interventions and the degree to which the objectives and nature of the evaluation of the intervention may differ as prototypes are developed and evaluators gain experience [57]. The unique role of eHealth interventions also needs to be considered; as these inherently deal with people's well-being, there is a need for closer evaluation of the psychological aspects of a user's experience.

Overall, the evaluation of users' psychological experiences with psychosocial interventions is a new yet growing area. Of the 21 studies in the final selection, only 10 were explicitly focused on understanding users' psychological experience (as opposed to usability, outcome, or overall evaluation). Of those 10 studies, none were published before 2009 and 6 were published in the past 2 years. The nascency of the topic and exploratory nature of this study means that a number of opportunities for future research and action in the area are possible.

Methodological Limitations

The methodological limitations of this study derived from the cross-disciplinary nature of the topic, resource constraints, and the newness of the field that could have resulted in missed studies. Searching across databases from the computer science and health and social science fields precluded the use of consistent search protocols in terms of qualifiers and data fields searched. Indeed, as noted by reviewers, several studies were missed in the search process [58-61]. Examining these studies, however, offered no additional methods of user inquiry. For example, the Helpful Aspects of Therapy questionnaire used by Gega et al [40] was also used by Richards and Timulak [61]. In addition, the search process was limited to scholarly material.

In hindsight, the resource constraint could have been mediated somewhat by limiting the search date to studies published in the past 5 years.

The retrieved studies were not subjected to a quality review. This was due to the nature of the study that was looking at methods used rather than study outcomes. Overall, one could then argue that the search was semisystematic with the front end (the search process) meeting the requirements of a predetermined and comprehensive search, but the back end (data extraction and synthesis) was less systematized and more exploratory in nature. It reflects the challenges outlined by Jesson et al [62] and Curran et al [30] in conducting systematic reviews across disciplines.

There is a lack of consistency in the terminology used to describe Internet-delivered interventions. An effort to create and broadly ratify agreed key or subject words would make the dissemination of information much easier. "eHealth" is almost as broad a term as "health"—searching for eHealth (or mHealth) interventions is as wide a scope as searching for health interventions. Therefore, as the area grows and consolidates, categorization needs to be clearer and differentiate from other deliveries such as telemedicine. Interventions need to be defined more by type (eg, CBT, mindfulness, psychoeducation) with an agreed identifier such as Web-based, Internet-based, or online for example.

Future Research and Implications

With regard to Web-based psychosocial interventions, there is a lack of cohesion between computer science literature, focused on the technical design, and health literature, focused on the treatment process and its impact. Combining modes of assessment prevalent in one discipline, such as usability testing from computer science, with session and treatment evaluations from psychology and social sciences could bridge this gap. This collaboration could contribute to the development of best practice protocols for understanding users' psychological experience that might include an evaluation of the in situ experience of the participant using the system, postsession impact, and a reflective posttreatment review. For example, combining the approach by Serowik et al [49] using a feelings-based think aloud usability and a posttreatment therapeutic evaluation (such as the Working Alliance Inventory or CCI) with the type of session evaluations used by Gega et al [40] would provide a comprehensive view of the user's experience, as would a user experience–focused usability inquiry method such as situated co-inquiry [54]. Research approaches focused on capturing user experience in situ help us understand the impact of Web-based interventions moment-to-moment, as well as their overall effectiveness. There is a role for clinicians in this process along with computer usability and design experts.

Overall, it is the experience of the user that is important in delivering acceptable and useful treatment. An understanding of user experience including their expectations and responses is essential to increasing acceptability, effectiveness, and adherence to cost-effective, broadly accessible Web-based psychosocial interventions. The therapeutic process is important to treatment and the way it is assessed needs to keep up with the way in which therapy is delivered. As the delivery of health changes, there needs to be increasing collaboration among disciplines. This will contribute not only to robust best practice but also to the creation of new and agreed terminology and a cohesive body of literature to ensure broad and effective dissemination of knowledge. A critical understanding of user experience of eHealth is needed to improve outcomes for people who look to the Internet for help.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Database search details.

[PDF File (Adobe PDF File), 31KB - jmir_v18i6e181_app1.pdf]

Multimedia Appendix 2

Included studies.

[PDF File (Adobe PDF File), 43KB - jmir_v18i6e181_app2.pdf]

Multimedia Appendix 3

Excluded studies.

[PDF File (Adobe PDF File), 45KB - jmir_v18i6e181_app3.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.   Al-Shorbaji N, Geissbuhler A. Establishing an evidence base for e-health: the proof is in the pudding. Bull World Health Organ 2012 May 1;90(5):322-322A [FREE Full text] [doi: 10.2471/BLT.12.106146] [Medline: 22589560]

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

4.   Muñoz RF. Using evidence-based internet interventions to reduce health disparities worldwide. J Med Internet Res 2010 Dec;12(5):e60 [FREE Full text] [doi: 10.2196/jmir.1463] [Medline: 21169162]

5.   Viswanath K, Kreuter MW. Health disparities, communication inequalities, and eHealth. Am J Prev Med 2007 May;32(5 Suppl):S131-S133 [FREE Full text] [doi: 10.1016/j.amepre.2007.02.012] [Medline: 17466818]

6.   Hesse BW, Shneiderman B. eHealth research from the user's perspective. Am J Prev Med 2007 May;32(5 Suppl):S97-103 [FREE Full text] [doi: 10.1016/j.amepre.2007.01.019] [Medline: 17466825]

7.   Andersson G, Paxling B, Wiwe M, Vernmark K, Felix CB, Lundborg L, et al. Therapeutic alliance in guided internet-delivered cognitive behavioural treatment of depression, generalized anxiety disorder and social anxiety disorder. Behav Res Ther 2012 Sep;50(9):544-550. [doi: 10.1016/j.brat.2012.05.003] [Medline: 22728647]

8.   Gainsbury S, Blaszczynski A. A systematic review of Internet-based therapy for the treatment of addictions. Clin Psychol Rev 2011 Apr;31(3):490-498. [doi: 10.1016/j.cpr.2010.11.007] [Medline: 21146272]

9.   Nordgren LB, Hedman E, Etienne J, Bodin J, Kadowaki A, Eriksson S, et al. Effectiveness and cost-effectiveness of individually tailored Internet-delivered cognitive behavior therapy for anxiety disorders in a primary care population: a randomized controlled trial. Behav Res Ther 2014 Aug;59:1-11 [FREE Full text] [doi: 10.1016/j.brat.2014.05.007] [Medline: 24933451]

10.   Health Promotion Agency. The Journal. URL: http://www.depression.org.nz [accessed 2015-10-27] [WebCite Cache ID 6caVGxby4]

11.   Sleepio. Let's build your sleep improvement program URL: https://www.sleepio.com/ [accessed 2015-10-27] [WebCite Cache ID 6caVt1au6]

12.   Barak A, Hen L, Boniel-Nissim M, Shapira N. A Comprehensive Review and a Meta-Analysis of the Effectiveness of Internet-Based Psychotherapeutic Interventions. Journal of Technology in Human Services 2008 Jul 03;26(2-4):109-160. [doi: 10.1080/15228830802094429]

13.   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]

14.   Black AD, Car J, Pagliari C, Anandan C, Cresswell K, Bokun T, et al. The impact of eHealth on the quality and safety of health care: a systematic overview. PLoS Med 2011 Jan;8(1):e1000387 [FREE Full text] [doi: 10.1371/journal.pmed.1000387] [Medline: 21267058]

15.   van Heerden A, Tomlinson M, Swartz L. Point of care in your pocket: a research agenda for the field of m-health. Bull World Health Organ 2012 May 1;90(5):393-394 [FREE Full text] [doi: 10.2471/BLT.11.099788] [Medline: 22589575]

16.   Buenestado D, Elorz J, Pérez-Yarza EG, Iruetaguena A, Segundo U, Barrena R, et al. Evaluating acceptance and user experience of a guideline-based clinical decision support system execution platform. J Med Syst 2013 Apr;37(2):9910. [doi: 10.1007/s10916-012-9910-7] [Medline: 23377779]

17.   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]

18.   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]

19.   Kring A, Johnson S, Davison G, Neale J. Abnormal Psychology. 12th edition. New York: John Wiley and Sons; 2012.

20.   Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. BMJ 2000 Sep 16;321(7262):694-696 [FREE Full text] [Medline: 10987780]

21.   Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]

22.   Hill C. Helping Skills: Facilitating Exploration, Insight, and Action. 3rd edition. Washington, DC: American Psychological Association (APA); 2009.

23.   Martin DJ, Garske JP, Davis MK. Relation of the therapeutic alliance with outcome and other variables: a meta-analytic review. J Consult Clin Psychol 2000 Jun;68(3):438-450. [Medline: 10883561]

24.   Wu C, Tzeng Y, Kuo B, Tzeng G. Integration of affective computing techniques and soft computing for developing a human affective recognition system for U-learning systems. IJMLO 2014;8(1):50-66. [doi: 10.1504/ijmlo.2014.059997]

25.   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]

26.   Luborsky L, Barber JP, Siqueland L, Johnson S, Najavits LM, Frank A, et al. The Revised Helping Alliance Questionnaire (HAq-II) : Psychometric Properties. J Psychother Pract Res 1996;5(3):260-271 [FREE Full text] [Medline: 22700294]

27.   Miller S, Duncan B, Brown J, Sparks J, Claud D. The outcome rating scale: A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. Journal of Brief Therapy 2003;2(2):91-100.

28.   Cooper A, Reimann R, Cronin D. About face 3: the essentials of interaction design. 3rd edition. Indianapolis, IN: Wiley Pub; 2007.

29.   Zhang Z. Usability evaluation. In: Zaphiris P, editor. Human Computer Interaction Research in Web Design and Evaluation. Hershey, PA: IGI Global; 2007:209-228.

30.   Curran C, Burchardt T, Knapp M, McDaid D, Bingqin L. Challenges in Multidisciplinary Systematic Reviewing: A Study on Social Exclusion and Mental Health Policy. Social Policy & Administration 2007;41(3):289-312. [doi: 10.1111/j.1467-9515.2007.00553.x]

31.   Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10 [FREE Full text] [doi: 10.1186/1471-2288-7-10] [Medline: 17302989]

32.   Knowles SE, Toms G, Sanders C, Bee P, Lovell K, Rennick-Egglestone S, et al. Qualitative meta-synthesis of user experience of computerised therapy for depression and anxiety. PLoS One 2014;9(1):e84323 [FREE Full text] [doi: 10.1371/journal.pone.0084323]

33.   Baños RM, Etchemendy E, Farfallini L, García-Palacios A, Quero S, Botella C. EARTH of Well-Being System: A pilot study of an Information and Communication Technology-based positive psychology intervention. The Journal of Positive Psychology 2014 Jun 23;9(6):482-488. [doi: 10.1080/17439760.2014.927906]

34.   Bendelin N, Hesser H, Dahl J, Carlbring P, Nelson KZ, Andersson G. Experiences of guided Internet-based cognitive-behavioural treatment for depression: a qualitative study. BMC Psychiatry 2011;11:107 [FREE Full text] [doi: 10.1186/1471-244X-11-107] [Medline: 21718523]

35.   Bradley KL, Robinson LM, Brannen CL. Adolescent help-seeking for psychological distress, depression, and anxiety using an Internet program. International Journal of Mental Health Promotion 2012 Feb;14(1):23-34. [doi: 10.1080/14623730.2012.665337]

36.   Cartreine JA, Locke SE, Buckey JC, Sandoval L, Hegel MT. Electronic problem-solving treatment: description and pilot study of an interactive media treatment for depression. JMIR Res Protoc 2012;1(2):e11 [FREE Full text] [doi: 10.2196/resprot.1925] [Medline: 23611902]

37.   de Graaf LE, Huibers MJ, Riper H, Gerhards SA, Arntz A. Use and acceptability of unsupported online computerized cognitive behavioral therapy for depression and associations with clinical outcome. J Affect Disord 2009 Aug;116(3):227-231. [doi: 10.1016/j.jad.2008.12.009] [Medline: 19167094]

38.   Devi R, Carpenter C, Powell J, Singh S. Exploring the experience of using a web-based cardiac rehabilitation programme in a primary care angina population: a qualitative study. International Journal of Therapy & Rehabilitation 2014;21(9):434-440. [doi: 10.12968/ijtr.2014.21.9.434]

39.   Fergus KD, McLeod D, Carter W, Warner E, Gardner SL, Granek L, et al. Development and pilot testing of an online intervention to support young couples' coping and adjustment to breast cancer. Eur J Cancer Care (Engl) 2014 Jul;23(4):481-492. [doi: 10.1111/ecc.12162] [Medline: 24472013]

40.   Gega L, Smith J, Reynolds S. Cognitive behaviour therapy (CBT) for depression by computer vs. therapist: patient experiences and therapeutic processes. Psychother Res 2013;23(2):218-231. [doi: 10.1080/10503307.2013.766941]

41.   Gerhards SA, Abma TA, Arntz A, de Graaf LE, Evers SM, Huibers MJ, et al. Improving adherence and effectiveness of computerised cognitive behavioural therapy without support for depression: a qualitative study on patient experiences. J Affect Disord 2011 Mar;129(1-3):117-125. [doi: 10.1016/j.jad.2010.09.012] [Medline: 20889214]

42.   Gorlick A, Bantum EO, Owen JE. Internet-based interventions for cancer-related distress: exploring the experiences of those whose needs are not met. Psychooncology 2014 Apr;23(4):452-458 [FREE Full text] [doi: 10.1002/pon.3443] [Medline: 24243756]

43.   Gulec H, Moessner M, Mezei A, Kohls E, Túry F, Bauer S. Internet-based maintenance treatment for patients with eating disorders. Professional Psychology: Research and Practice 2011;42(6):479-486. [doi: 10.1037/a0025806]

44.   Hind D, O'Cathain A, Cooper CL, Parry GD, Isaac CL, Rose A, et al. The acceptability of computerised cognitive behavioural therapy for the treatment of depression in people with chronic physical disease: a qualitative study of people with multiple sclerosis. Psychol Health 2010 Jul;25(6):699-712. [doi: 10.1080/08870440902842739] [Medline: 20204955]

45.   Lara M, Tiburcio M, Aguilar Abrego A, Sánchez-Solís A. A four-year experience with a Web-based self-help intervention for depressive symptoms in Mexico. Rev Panam Salud Publica 2014;35(5-6):399-406 [FREE Full text] [Medline: 25211568]

46.   Lederman R, Wadley G, Gleeson J, Bendall S, Álvarez-Jiménez M. Moderated online social therapy. ACM Trans. Comput.-Hum. Interact 2014 Feb 01;21(1):1-26. [doi: 10.1145/2513179]

47.   Lillevoll KR, Wilhelmsen M, Kolstrup N, Høifødt RS, Waterloo K, Eisemann M, et al. Patients' experiences of helpfulness in guided internet-based treatment for depression: qualitative study of integrated therapeutic dimensions. J Med Internet Res 2013 Jun;15(6):e126 [FREE Full text] [doi: 10.2196/jmir.2531] [Medline: 23786763]

48.   McClay C, Waters L, McHale C, Schmidt U, Williams C. Online cognitive behavioral therapy for bulimic type disorders, delivered in the community by a nonclinician: qualitative study. J Med Internet Res 2013 Mar;15(3):e46 [FREE Full text] [doi: 10.2196/jmir.2083] [Medline: 23502689]

49.   Serowik K, Ablondi K, Black A, Rosen M. Developing a benefits counseling website for Veterans using Motivational Interviewing techniques. Computers in Human Behavior 2014 Aug;37:26-30. [doi: 10.1016/j.chb.2014.03.019]

50. Tonkin-Crine S, Bishop FL, Ellis M, Moss-Morris R, Everitt H. Exploring patients' views of a cognitive behavioral therapy-based website for the self-management of irritable bowel syndrome symptoms. J Med Internet Res 2013 Sep;15(9):e190 [FREE Full text] [doi: 10.2196/jmir.2672] [Medline: 24001787]

51. Topolovec-Vranic J, Cullen N, Michalak A, Ouchterlony D, Bhalerao S, Masanic C, et al. Evaluation of an online cognitive behavioural therapy program by patients with traumatic brain injury and depression. Brain Inj 2010 Apr;24(5):762-772. [doi: 10.3109/02699051003709599] [Medline: 20370383]

52. Van Voorhees BW, Ellis JM, Gollan JK, Bell CC, Stuart SS, Fogel J, et al. Development and process evaluation of a primary care internet-based intervention to prevent depression in emerging adults. Prim Care Companion J Clin Psychiatry 2007;9(5):346-355 [FREE Full text] [Medline: 17998953]

53. Wade SL, Oberjohn K, Burkhardt A, Greenberg I. Feasibility and preliminary efficacy of a web-based parenting skills program for young children with traumatic brain injury. J Head Trauma Rehabil 2009;24(4):239-247. [doi: 10.1097/HTR.0b013e3181ad6680] [Medline: 19625863]

54. Carter P. Liberating usability testing. Interactions 2007 Mar 01;14(2):18-22. [doi: 10.1145/1229863.1229864]

55. Myers MD, Newman M. The qualitative interview in IS research: Examining the craft. Information and Organization 2007 Jan;17(1):2-26. [doi: 10.1016/j.infoandorg.2006.11.001]

56. Goffman E. The presentation of self in everyday life. New York: Doubleday; 1959.

57. Greenhalgh T, Russell J. Why do evaluations of eHealth programs fail? An alternative set of guiding principles. PLoS Med 2010;7(11):e1000360 [FREE Full text] [doi: 10.1371/journal.pmed.1000360] [Medline: 21072245]

58. Advocat J, Lindsay J. Internet-based trials and the creation of health consumers. Soc Sci Med 2010 Feb;70(3):485-492. [doi: 10.1016/j.socscimed.2009.10.051] [Medline: 19926185]

59. Iloabachie C, Wells C, Goodwin B, Baldwin M, Vanderplough-Booth K, Gladstone T, et al. Adolescent and parent experiences with a primary care/Internet-based depression prevention intervention (CATCH-IT). General Hospital Psychiatry 2011 Nov;33(6):543-555. [doi: 10.1016/j.genhosppsych.2011.08.004] [Medline: 21958447]

60. Mitchell N, Dunn K. Pragmatic evaluation of the viability of CCBT self-help for depression in higher education. Counselling and Psychotherapy Research 2007 Sep;7(3):144-150. [doi: 10.1080/14733140701565987]

61. Richards D, Timulak L. Client-identified helpful and hindering events in therapist-delivered vs. self-administered online cognitive-behavioural treatments for depression in college students. Counselling Psychology Quarterly 2012 Sep;25(3):251-262. [doi: 10.1080/09515070.2012.703129]

62. Jesson J, Matheson L, Lacey FM. Doing your literature review: Traditional and systematic techniques. London: SAGE Publications Ltd; 2011.

Abbreviations

AMSTAR: Assessment of Multiple SysTemAtic Reviews
CBT: cognitive behavioral therapy
CCI: Client Change Interview
RCT: randomized clinical trial
TBI: traumatic brain injury

XSL•FO
RenderX

Original Paper

Feasibility and Performance Test of a Real-Time Sensor-Informed Context-Sensitive Ecological Momentary Assessment to Capture Physical Activity

Genevieve Fridlund Dunton[1], PhD, MPH; Eldin Dzubur[1], MS; Stephen Intille[2], PhD

[1]Department of Preventive Medicine, University of Southern California, Los Angeles, CA, United States

[2]College of Computer and Information Science & Dept. of Health Sciences, Bouvé College of Health Sciences, Northeastern University, Boston, MA, United States

Corresponding Author:
Genevieve Fridlund Dunton, PhD, MPH
Department of Preventive Medicine
University of Southern California
2001 N Soto St.
Los Angeles, CA,
United States
Phone: 1 3238650805
Fax: 1 3238650125
Email: dunton@usc.edu

Abstract

Background: Objective physical activity monitors (eg, accelerometers) have high rates of nonwear and do not provide contextual information about behavior.

Objective: This study tested performance and value of a mobile phone app that combined objective and real-time self-report methods to measure physical activity using sensor-informed context-sensitive ecological momentary assessment (CS-EMA).

Methods: The app was programmed to prompt CS-EMA surveys immediately after 3 types of events detected by the mobile phone's built-in motion sensor: (1) Activity (ie, mobile phone movement), (2) No-Activity (ie, mobile phone nonmovement), and (3) No-Data (ie, mobile phone or app powered off). In addition, the app triggered random (ie, signal-contingent) ecological momentary assessment (R-EMA) prompts (up to 7 per day). A sample of 39 ethnically diverse high school students in the United States (aged 14-18, 54% female) tested the app over 14 continuous days during nonschool time. Both CS-EMA and R-EMA prompts assessed activity type (eg, reading or doing homework, eating or drinking, sports or exercising) and contextual characteristics of the activity (eg, location, social company, purpose). Activity was also measured with a waist-worn Actigraph accelerometer.

Results: The average CS-EMA + R-EMA prompt compliance and survey completion rates were 80.5% and 98.5%, respectively. More moderate-to-vigorous intensity physical activity was recorded by the waist-worn accelerometer in the 30 minutes before CS-EMA activity prompts (M=5.84 minutes) than CS-EMA No-Activity (M=1.11 minutes) and CS-EMA No-Data (M=0.76 minute) prompts (P's<.001). Participants were almost 5 times as likely to report going somewhere (ie, active or motorized transit) in the 30 minutes before CS-EMA Activity than R-EMA prompts (odds ratio=4.91, 95% confidence interval=2.16-11.12).

Conclusions: Mobile phone apps using motion sensor–informed CS-EMA are acceptable among high school students and may be used to augment objective physical activity data collected from traditional waist-worn accelerometers.

KEYWORDS

mobile phones; ecological momentary assessment; accelerometer; physical activity

Introduction

Surveillance, epidemiological, and intervention studies typically use either retrospective self-report measures or objective sensor-based monitors (eg, accelerometers, heart rate devices) to capture physical activity behavior in youth [1-3]. Retrospective self-report methods ask participants to recall levels and types of activities performed over the past few days, weeks, or year [4]. As such, they can be vulnerable to various types of reporting errors and biases [5]. Although objective activity monitors may yield more valid measures of physical activity [6] and are being deployed in large-scale surveillance studies with children and adolescents [1], they may result in significant amounts of missing data due to device nonwear [7,8]. Participants forget to wear or carry monitors, and they remove monitors when they do not want to or cannot wear them. Often, these data are not missing at random because youth remove the monitors during physical activity bouts such as swimming or those bouts involving physical contact with others (eg, football, soccer), which can result in biased activity estimates [9].

In addition to high rates of device nonwear, objective activity monitors are unable to capture descriptive and contextual information about physical activity and sedentary behavior. Accelerometers and other objective activity monitors typically assess the intensity and duration of physical body movement [10,11]. Yet, most devices cannot measure the type and purpose of the activity, emotional responses to the activity, where and with whom the activity occurred, and other cognitive and motivational factors. Information about these real-time psychological and environmental correlates of physical activity has growing importance for the development of just-in-time adaptive interventions to encourage physical activity at the point of decision-making during adolescents' everyday lives [12,13].

Methods of physical activity assessment that collect real-time self-report data about activities and contexts during key moments of the day, such as when an accelerometer is removed, or immediately after a bout of physical activity, have the potential to yield information that an accelerometer cannot. Ecological momentary assessment (EMA) is one such real-time self-report data capture method, which elicits responses to electronic surveys throughout the course of daily life [14,15]. EMA has many advantages over retrospective self-report measures such as reducing memory errors and biases and collecting more ecologically valid assessments in one's natural environment [16-18]. Standard EMA uses *interval-contingent* sampling to trigger surveys at preset times (eg, at 8 am and 12 noon everyday), *signal-contingent* sampling to trigger surveys at random times throughout the day, or *event-contingent* sampling to trigger surveys during or after a predetermined behavior such as a bout of physical activity [16]. However, these sampling strategies suffer from a number of limitations. Interval- and signal-contingent sampling strategies often fail to capture less common behavior, such as physical activity, as it is occurring. Event-contingent sampling requires the participant to self-initiate surveys during or after particular behavior, which is prone to delayed reports and the failure to report events altogether. CS-EMA is an innovative strategy that addresses these problems by automatically triggering survey prompts at opportune times based on detected information from internal or external sensors to collect real-time information about key behaviors, events, or contexts [19].

This study tested a mobile phone app called Mobile Teen that implemented CS-EMA by performing real-time analysis on data gathered about the mobile phone's movement from the mobile phone's built-in motion sensor. The app detected major transitions in type of mobile phone movement and subsequently triggered real-time electronic surveys that collected self-report information about recent physical activity and sedentary behavior. A growing number of studies have measured physical activity using native mobile features with varying levels of accuracy [20,21]. The goals of this study were to examine the feasibility, performance, accuracy, and utility of the Mobile Teen EMA app for capturing information about physical activity behaviors and contexts in adolescents. Specially, the objectives were to: (1) describe EMA survey compliance, response latency, and completion time and rates, in addition to how the mobile phone was carried; (2) evaluate the performance of the CS-EMA prompting features by examining differences in reported and objectively measured activity levels across the different EMA prompt types; (3) describe the extent to which EMA provides data about activity during periods of waist-worn accelerometer nonwear; and (4) describe contextual characteristics of EMA-reported sports and exercise episodes.

Methods

Mobile Teen App

The Mobile Teen app was designed for mobile phones running the Android operating system. The software was written in Java and targeted Android versions 2.3.3 to 4.3, the versions available at the time of the research. The Mobile Teen app was programmed to conduct sensor-informed CS-EMA by using the mobile phone's built-in motion sensor to automatically detect periods of (1) Activity (15+ minutes of high-intensity activity followed by 10+ minutes of low-intensity activity); (2) No-Activity (60+ minutes of low-intensity activity followed by 2+ minutes of moderate-intensity activity); and (3) No-Data (10+ minutes of no activity data followed by 1+ minutes of some activity data). The app then used these sensor-informed movement transition cues to trigger real-time CS-EMA self-report surveys measuring the type and purpose of activity previously performed, enjoyment of that activity, and social and physical features of the activity setting. A more detailed description of the design and development of the Mobile Teen app is available elsewhere [22].

Participants and Recruitment

The study sample consisted of low-to-middle–income adolescents in grades 9 to 12. Participants were recruited through an urban Los Angeles area high school using informational fliers, posters, and classroom visits. Inclusion criteria were as follows: (1) student in grade 9 to 12 enrolled at the participating high school, (2) able to comprehend written English, (3) no health or physical limitations that prohibit regular physical activity, and (4) regular use of an Android or global system for mobile communication (GSM)-based mobile phone with service provided by AT&T or T-Mobile on a standard mobile phone

contract. For adolescents who did not own an Android mobile phone, the restriction to having a GSM-based mobile provider (AT&T or T-Mobile) was made so that their personal phone subscriber identity module (SIM) cards could be easily switched to temporary LG Nexus 4 mobile phones with the Mobile Teen app installed for the duration of the study. Doing so allowed participants to use the study mobile phone to make and receive calls and short message service (SMS) messages through personal mobile phone numbers. Adolescents who expressed interest in the study during a school visit were called by phone to be screened for eligibility and scheduled for an orientation session with research staff. Parental consent and child assent were obtained. Study procedures were approved by the Institutional Review Boards at the University of Southern California and Northeastern University.

Procedures

Participants attended a 45-minute orientation session at the high school on a weekday afternoon. During this session, they completed a questionnaire and anthropometric assessments. Research staff assisted participants with either downloading the Mobile Teen app to their personal Android mobile phone or moving their SIM card from a personal mobile phone into a loaned Android study mobile phone with the Mobile Teen app preinstalled. They were also given instructions on how to complete EMA surveys. Over the next 28 days, participants were asked to wear the accelerometer and proceed with their daily routines as normal. Participants were asked to carry the mobile phone as usual (in their pockets, hands, purses, or bags) during waking hours. On 14 of those days (either the first 2 weeks or the second 2 weeks, as randomly assigned), they received EMA survey prompts several times per day during nonschool waking hours (3-9 pm on weekdays and 7 am-9 pm on weekend days). Research staff members contacted participants twice by phone during the 28-day monitoring period to encourage compliance and address technical issues. Participants received up to $180 for completing the study.

Measures

Ecological Momentary Assessment

The Mobile Teen app triggered mobile phone sensor-informed CS-EMA prompts according to the 3 prompting rules described previously. In addition, the app triggered random (ie, signal-contingent) R-EMA prompts up to 3 times per day on weekdays and 7 times per day on weekend days. On receiving an auditory EMA prompt (a pleasant, but loud, 4-second chime), participants were instructed to stop their current activity and complete a short electronic survey question sequence using the touch screen of the mobile phone. Participants were allowed to set mobile phones to mute or vibrate to prevent interruption. If an EMA prompt occurred during an incompatible activity (eg, sleeping, bathing, or driving), participants were instructed to

ignore it. If no entry was made, the app emitted up to 2 reminder prompts at 3-minute intervals. After this point, the electronic EMA survey became inaccessible until the next prompting opportunity. To avoid excessive prompting, the app enforced a 30-minute gap between all prompts.

EMA items assessed major activity types, methods of carrying the mobile phone, reasons for not carrying the mobile phone, and other psychological and contextual factors related to behavior (See Figure 1). CS-EMA prompts asked, "What have you been doing between (start time) and (stop time)?" with times inserted by the app based on information from the built-in mobile phone motion sensor. Alternatively, the R-EMA question sequence began with the *activity type* question, "What have you been doing in the last 30 minutes?" For both the CS-EMA and R-EMA activity-type questions, a response structure was used where participants could indicate 1 or more activities (ie, "choose all that apply") that were performed during the time period. Response options include, "*Reading or doing homework; Using technology (TV, phone); Eating/Drinking; Sports/Exercising; Going somewhere; Hanging out; Other.*" If *Other* was selected, an extra question listed additional options including, "*Doing chores/Cooking; Showering/Bathing; Sleeping; Working/Part-time job; Getting ready for something; Shopping; Getting dressed; Class/School; Playing with children; Playing catch; Waiting; Doing something else.*" If *Sports/Exercising* was reported as an activity type, a follow-up question asked about the specific type of sports or exercise activity performed (eg, *Basketball/Football/Soccer, Bicycling, Other Running/Jogging, Exercise/Dance/Karate class, Weightlifting/Strength training).*

For each type of activity reported, participants were asked to indicate how long it lasted (in minutes), the body position (eg, *Lying down, Sitting, Standing*), how the mobile phone was carried (eg, *On my belt, In my pocket, Not with me*) and the reason for not having the mobile phone with them (eg, *Forgot it, Did not want to damage it, Too uncomfortable*). If *Sports/Exercising* was reported as an activity type, a unique question branching sequence was initiated that asked about type of fitness skill involved (eg, *Flexibility, Strengthening, Endurance*), extra weight carried (eg, *None, Less than 5 lbs, 5-10 lbs*), degree of incline (eg, *Mainly going uphill, Mainly going downhill, Mainly staying on flat ground*), the physical context (eg, *Home, Work, School*), the main purpose (eg, *Fun/Recreation; To get somewhere; For work, homework, or housework*), how enjoyable it was (eg, *Not at all, A little, Moderately*), and the social context (eg, *Alone or With Friends, Parents, Siblings*). The branching sequence pertaining to sports and exercise activities was only initiated in a randomly programmed 40% of applicable surveys as a method of limiting potential subject response burden.

Figure 1. Mobile Teen app EMA screenshots, which include 8 sample survey items from the Mobile Teen app's EMA interface. The items in the top row query (left to right) activity type during random EMA prompts, activity type during a context-sensitive EMA prompt, how the phone was carried, and reasons for not carrying the device (if it was not carried). Items on the bottom row (left to right) query specific type of sports or exercise, fitness skill involved, physical context, and social context.
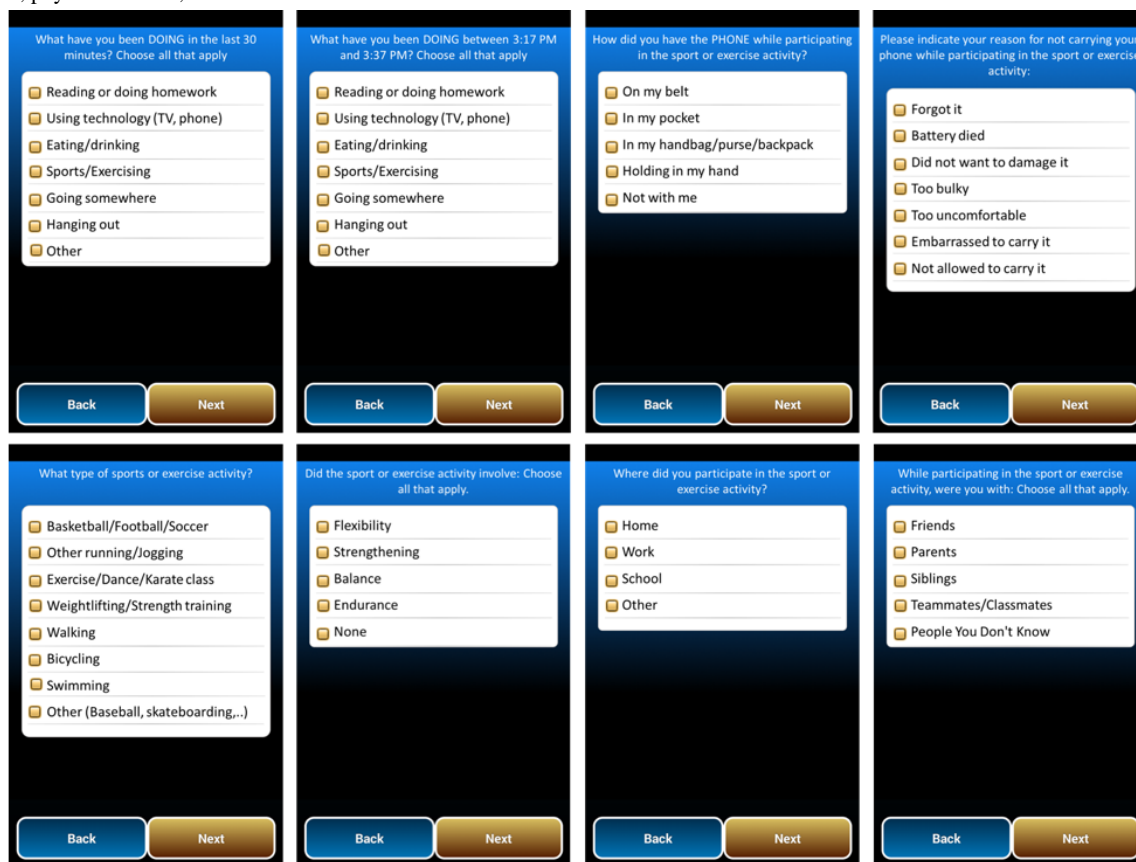
Waist-Worn Accelerometer

The Actigraph, Inc, activity monitor provided an objective measure of physical activity. A combination of GT1M, GT2M, and GT3X models were used. The device was worn on the right hip attached to an adjustable belt. Activity monitors were not worn while sleeping, bathing, or swimming. A 30-second epoch was used. All accelerometer recordings were time-stamped to be linked with EMA data captured on the mobile phone. Outcome variables consisted of the total number of minutes of moderate-to-vigorous physical activity (MVPA) occurring within the 30-minute window before each EMA prompt. MVPA was defined using age-specific thresholds generated from the Freedson prediction equation (≥ 4 metabolic equivalents [METs]) [23]. Strings of continuous zero activity counts lasting 60 minutes or more were considered to be periods of waist accelerometer nonwear (with an allowance of up to 2 minutes with nonzero activity counts during that period) [1].

Body Mass Index

Research staff measured height and weight using an electronically calibrated digital scale (Tanita WB-110A) and professional stadiometer (PE-AIM-101) to the nearest 0.1 kg and 0.1 cm, respectively. Body mass index (BMI) was calculated (kg/m^2) and converted to Centers for Disease Control and Prevention age- and gender-specific BMI percentiles.

Demographic

Participants' age, gender, year in school, and ethnicity were assessed through a self-report paper-and-pencil questionnaire. Parents reported annual household income.

Data Processing and Analyses

Descriptive statistics for demographic characteristics were calculated with individual participant as the unit of analysis (Level-2); all other descriptive statistics used occasions (ie, EMA prompts) as the unit of analysis (Level-1). Data were analyzed with multilevel modeling in Stata (version 14). Multilevel models adjust the standard errors for clustering of EMA prompts (Level-1) within people (Level-2) [24]. Between-subject (BS) and within-subject (WS) versions (ie, partitioning the variance) of the main effects were generated. The BS version represents the individual mean deviation from the grand mean, and the WS version represents deviation from one's own mean at any given prompt [25]. Similarly, the BS and WS variation terms for binary predictors were created using grand mean-centering (ie, subtracting by the group mean proportion) and person mean-centering (ie, subtracting by the individual mean proportion) methods, respectively [26].

Descriptive statistics (ie, mean, standard deviations [SDs], frequencies, and percentages) were calculated to examine EMA compliance, response latency, and completion time and rates, as well as how the mobile phone was carried (objective #1). In addition, a series of multilevel logistic regressions were run to

examine whether any demographic or temporal variables were associated with EMA compliance (binary outcome categorized as 0=unanswered prompt and 1=answered prompt), methods of carrying the mobile phone (binary outcome categorized as 1=not with me and 0=all else), and reason for not having the mobile phone with him or her (binary outcome categorized as 1=did not want to damage it and 0=all else).

Multilevel linear regression models were computed to compare EMA prompt types in terms of self-reported and objectively measured activity levels (objective#2). The models tested whether there were differences across the 4 EMA prompt types (ie, R-EMA, CS-EMA Activity, CS-EMA No-Activity, and CS-EMA No-Data) in terms of objective levels of MVPA (measured by the waist-worn Actigraph) in the 30 minutes before the prompt. In addition, multilevel logistic linear regression models tested whether there were differences across prompt types (ie, R-EMA, CS-EMA Activity, CS-EMA No-Activity, and CS-EMA No-Data) in terms of the likelihood of engaging in each type of activity (vs all other activities) in the time leading up to the EMA prompt (ie, 30-minute intervals before R-EMA prompts and automatically detected time intervals before CS-EMA prompts). Each type of activity was tested in a separate logistic regression model. The 4-level EMA prompt type served as the independent variable (with R-EMA designated as the reference group) for the linear and logistic regressions testing the second study objective. Models testing the likelihood of engaging in each type of activity across EMA prompt types and models examining objective MVPA across prompt types were also adjusted by the day of week, time of day, chronological day in study, age, and gender. Random intercept models were estimated. Descriptive statistics (ie, means, SDs, frequencies, and percentages) were run to describe the extent to which EMA provides data about activity during periods of waist-worn accelerometer nonwear (objective #3) and measures contextual characteristics of EMA-reported sports and exercise episodes (objective #4).

Results

Participant Characteristics

Of 248 participants who initially expressed interest in the study, 5 participants lost interest, 2 of them were excluded based on asthma history, 10 of them did not have mobile phones, 126 of them could not be reached for scheduling, and 37 of them canceled their scheduled appointments; the remaining 68 participants met the inclusion criteria and verbally agreed to join at a mobile phone screening interview. A total of 44 participants subsequently attended the scheduled orientation appointment with their parents and consented to participate. During the study period, 4 participants voluntarily dropped out, leaving a total of 40 participants who completed the study. One participant lost the waist-worn accelerometer, resulting in 39 participants with complete data for analysis. The analytic sample consisted of 21 girls and 18 boys between the ages of 14 and 18 (M=15.9 SD=1.2). Sixty-four percent reported Hispanic or Latino ethnicity, and all participants reported receiving free or reduced lunch at school. Table 1 summarizes further demographic characteristics of the sample.

Table 1. Demographic characteristics of participants (N=39).

Characteristics		N (%)
Gender		
	Male	18 (46.2)
	Female	21 (53.9)
BMI[a] (mean, SD[d])		24.58 (4.3)
Age (mean, SD)		15.90 (1.2)
Grade in school[b]		
	9	9 (23.7)
	10	11 (29.0)
	11	9 (23.7)
	12	9 (23.7)
Ethnicity[c]		
	White	7 (18.4)
	Black or African-American	1 (2.6)
	Hispanic/Latino	25 (65.8)
	Asian	4 (10.5)
	Other	1 (2.6)
Mobile phone platform		
	Google Android	27 (69.2)
	Apple iOS	12 (30.8)
Mobile phone device used	Loaned mobile phone	22 (56.4)
	Personal mobile phone	17 (43.6)

[a]BMI: body mass index.

[b]Information on grade in school was not reported for 1 participant.

[c]Ethnicity was not reported for 1 participant.

[d]SD: standard deviation.

Table 2. Ecological momentary assessment (EMA) prompts and compliance.

Variable	Total prompts	Average Compliance (%)[a]	SD[b] (%)	Range (%)
All prompts	3,631	80.5	14.6	38.2-97.0
R-EMA[c,g]	2,830	79.4	14.3	35.6-97.6
CS-EMA[h] activity[d]	40	87.8	28.5	0.0-100.0
CS-EMA No-Activity[e]	369	88.6	14.7	40.0-100.0
CS-EMA No-Data[f]	392	85.0	21.6	16.7-100.0

[a]Compliance (%) represents the number of EMA surveys answered divided by the number of EMA surveys prompted.

[b]SD: standard deviation

[c]Prompts occurring randomly within schedule intervals throughout the day.

[d]Prompts occurring after a 15-minute period of high-intensity activity followed by a 10-minute period of low-intensity activity.

[e]Prompts occurring after a 60-minute period of low-intensity activity followed by 2+ minutes of moderate-intensity activity.

[f]Prompts occurring after a 10-minute period of no activity data followed by a period of 1 minute of available data.

[g]R-EMA: random ecological momentary assessment.

[h]CS-EMA: context-sensitive ecological momentary assessment

XSL•FO
RenderX

EMA Compliance

Descriptive statistics for R-EMA and CS-EMA compliance rates are summarized in Table 2. Compliance rate was defined as the number of answered EMA surveys divided by the number of EMA surveys prompted when the mobile phone was powered on and charged. Prompts consisted of audible or tactile feedback, unless the app was in silent mode, in which case, the app would appear on the screen if the mobile phone was unlocked, but it did not otherwise alert the participant. In total, across all the participants, CS-EMA No-Activity and CS-EMA No-Data prompts were more common than CS-EMA Activity prompts. Average compliance rates were higher for CS-EMA prompts (84.8%, standard error of the mean [SE]=2.2%) than R-EMA prompts (78.8%, SE=2.4%; $z=3.67$, $P<.001$). In addition, average EMA compliance was higher when children were wearing a waist accelerometer in the 30 minutes leading up to the EMA prompt (83.0%, SE=1.9%) than when not wearing the waist accelerometer during this period (75.5%, SE=2.6%; $z=4.71$, $P<.001$). Average compliance to EMA prompts generally decreased across the study period. Individuals were 4.7% less likely to respond to an EMA prompt on a given day in the study compared with the day before, ($z=-4.8$, $P<.001$). However, compliance increased across the day. Participants were 8.7% more likely to respond to an EMA prompt at any given hour in a day compared with the hour before, ($z=5.32$, $P<.001$). Participants identifying as Hispanic were 2.21 times as likely to respond to any given EMA prompt than those identifying as any other ethnicity ($z=2.65$, $P<.01$). EMA compliance rates were unrelated to age, gender, BMI, grade in school, and household income (P values > .05). In total, 784 R-EMA surveys (25.8%) were not prompted at all either due to the mobile phone being powered off at the time the EMA prompt was scheduled or due to unknown technical problems. Unprompted R-EMA surveys were not included in the calculation of compliance rates.

EMA Response Latency and Completion Time and Rate

Overall, 2,862 R-EMA + CS-EMA surveys were completed (ie, all questions answered) out of the 2,907 survey prompts that were answered, yielding an EMA survey completion rate (once started) of 98.5%. A total of 2,372 (82.3%) completed EMA surveys were responded to after the first prompt, 318 surveys (11.0%) after the first reprompt (3 minutes later), and 193 surveys (6.7%) after the second reprompt (6 minutes later).

EMA survey completion time was defined as the time lag between receiving an EMA prompt and finishing the last question on the survey for that EMA prompt. EMA prompts requiring one or more reprompts (n=511) were not included in the calculation of completion time. The app was designed to time out if a survey was not completed within 15 minutes after the prompt. A total of 22 surveys that timed out for this reason were not included in the analysis of survey completion time. On average, surveys that were completed without reprompting were completed in 53.2 seconds (SD=47.27, range: 8-408). There were no differences in EMA completion time by age, gender, ethnicity, day of week, and time of day. However, participants completed surveys 1.61 seconds faster (confidence interval [CI]: 1.14-2.07) per chronological day in the study, ($z=-6.79$, $P<.001$; i.e., 20 seconds faster by the end of the study).

Methods of Carrying the Mobile Phone

Analyses examined how the mobile phone was carried by participants because movement of the mobile phone was designed to trigger CS-EMA prompting. Participants were instructed to carry the mobile phones as they normally would. That is, they were not asked to carry the mobile phone in a special way (eg, on the hip), to investigate how normal mobile phone usage could drive the sensor-informed CS-EMA feature of the Mobile Teen app even if the mobile phone is carried in a pocket or bag. Although a sensor not on the body would not capture body motion directly, large changes in motion (or lack of motion) could align with large changes in activity patterns. EMA prompts asked how the mobile phone was carried during each type of activity reported. When participants reported engaging in 2 or more activities (n=476) during any given EMA prompt, only one of those activities was randomly selected for inclusion in analysis of how the mobile phone was carried. Table 3 summarizes the distribution of methods for carrying the mobile phone device by gender. As compared with boys, girls were more than 3 times as likely to have the mobile phone *Within reach* (ie, close by but not physically on their bodies) and less likely to carry the mobile phone in their pockets. Across boys and girls, reasons provided for not having the mobile phone (ie, *Not with me*) during some portion of the window of time on which they were reporting activities were forgetting it elsewhere (15.4%), the battery died (10.3%), not wanting to damage it (29.1%), feeling it was too bulky (6.8%), feeling it was too uncomfortable (9.4%), being too embarrassed to carry it (5.1%), and not being allowed to carry it (23.9%).

Table 3. Methods of carrying the mobile phone device by gender.

Method of carrying device[a]	Boys (%)	Girls (%)	OR[b] (95% CI[c])[d]
Carrying in pocket	47.0	17.0	0.15 (0.06-0.35)
Holding in hand	18.4	25.6	1.72 (0.70-4.18)
Carrying in bag or purse	3.5	6.8	2.12 (0.82-5.51)
On belt	1.1	1.2	1.15 (0.57-2.35)
Within reach	21.7	41.6	3.11 (1.54-6.25)
Not with me	5.2	8.8	1.82 (0.88-3.77)

[a]Dependent variable was coded as 1=method of carrying device, 0=all other methods.

[b]OR: odds ratio

[c]CI: confidence interval.

[d]Gender was coded as 1=female, 0=male. Results were generated from multilevel logistic linear regression models testing whether there were gender differences across methods of carrying the mobile phone immediately before the prompt. Each method of carrying the mobile phone was tested in a separate multilevel logistic regression model. Percentages (%) represent the adjusted margins generated from the statistical models.

Analyses further examined differences in reasons for not having the mobile phone (ie, *Not with me*) by time of day, day in the study, and other demographic factors. For each additional hour in a day, participants were twice as likely to report not having the phone because the battery died (odds ratio [OR]=2.04; 95% CI=1.06-3.93), and less likely to report not having the mobile phone because the mobile phone was too bulky compared with any other reason (OR=0.75; 95% CI=0.58-0.98). Furthermore, participants were less likely to report not being allowed to have the mobile phone, compared with any other reason, for each additional chronological day in the study (OR=0.62; 95% CI=0.43-0.89). Gender, age, grade in school, ethnicity, income, BMI, and day of the week were not related to reported reasons for participants not having the mobile phone (*P* values>.05). Participants were more likely to report having the mobile phone before answered CS-EMA Activity (92.2%) than CS-EMA No-Activity (64.8%; z=−2.70, *P*<.01) or CS-EMA No-Data (67.3%; z=−3.05, *P*<.01) type EMA prompts.

Table 4. Type of activity reported by ecological momentary assessment (EMA) prompt type.

Activity type[a]	R-EMA[b, c]	CS-EMA[d] Activity[e]		CS-EMA No-Activity[f]		CS-EMA No-Data[g]	
	%	%	OR[i](95% CI[j])	%	OR (95% CI)	%	OR (95% CI)
Reading or doing homework	18.3	0.0	—	13.1	0.63 (0.44-0.90)	15.4	0.78 (0.50-1.22)
Using technology	24.6	2.9	0.08 (0.02-0.35)	29.2	1.30 (0.99-1.70)	16.9	0.59 (0.36-0.97)
Eating or drinking	10.0	19.3	2.18 (0.75-6.31)	8.3	0.81 (0.53-1.24)	9.9	0.99 (0.56-1.75)
Sports or exercise	3.7	10.2	3.66 (0.67-19.81)	7.8	2.54 (1.46-4.42)	4.6	1.30 (0.63-2.71)
Going somewhere	14.6	43.2	4.91 (2.16-11.13)	8.1	0.51 (0.33-0.78)	11.7	0.76 (0.47-1.24)
Hanging out	14.4	24.3	2.07 (0.90-4.76)	13.1	0.89 (0.61-1.30)	14.9	1.05 (0.66-1.67)
Sleeping	4.9	0	—	5.5	1.15 (0.64-2.07)	9.6	2.37 (1.29-4.35)
Other[h]	10.5	13.3	1.34 (0.50-3.57)	14.8	1.53 (1.07-2.20)	13.7	1.39 (0.87-2.22)

[a]Type of activity taking place during the time leading up to the EMA prompt (ie, 30-minute intervals before R-EMA prompts and automatically detected time intervals before CS-EMA prompts). Values in the table represent the proportion of the prompt type (column) reported in each activity type.

[b]R-EMA: random ecological momentary assessment.

[c]Prompts occurring randomly within scheduled intervals throughout the day.

[d]CS-EMA: context-sensitive ecological momentary assessment.

[e]Prompts occurring after a 15-minute period of high-intensity activity followed by a 10-minute period of low-intensity activity.

[f]Prompts occurring after a 60-minute period of low-intensity activity followed by 2+ minutes of moderate-intensity activity.

[g]Prompts occurring after a 10-minute period of no activity data followed by a period of 1 minute of available data.

[h]All other activities were recoded into the "Other" category. Results were generated from multilevel logistic linear regression models testing whether there were differences across prompt types (ie, R-EMA, CS-EMA Activity, CS-EMA No-Activity, and CS-EMA No-Data) in terms of the likelihood of engaging in each type of activity (vs all other activities) immediately before the prompt. Each type of activity was tested in a separate multilevel logistic regression model. EMA prompt type served as the independent variable with R-EMA designated as the reference group. Models were adjusted by day of week, time of day, chronological day in study, age, and gender. OR for WS effects are shown. There were no significant BS effects for any activity type (P values>.05; results not shown). Percentages (%) represent the adjusted margins generated from multiple independent statistical models controlling for the covariates, and therefore, the sum of a column may exceed 100% or fail to add up to 100%.

[i]OR: odds ratio.

[j]CI: confidence interval.

Differences in Activity Levels by EMA Prompt Types

Results indicate that there were significantly more MVPA minutes recorded in the 30-minute window before CS-EMA Activity prompts (M=6.02, SE=0.74 minutes) than CS-EMA No-Activity (M=1.12, SE=0.48 minutes; z=−7.94, P<.001) and CS-EMA No-Data (M=0.77, SE=0.48 minutes; z=−7.68, P<.001) prompts. Table 4 compares EMA prompt types in terms of the type of activity reported to take place in the time leading up to the EMA prompt (ie, 30-minute interval before R-EMA prompts and automatically detected time intervals before CS-EMA prompts). When participants reported engaging in 2 or more activities (n=476) during any given EMA prompt, only one of those activities was randomly selected for inclusion in this analysis. Participants were almost half as likely to report reading or doing homework before CS-EMA No-Activity prompts compared to R-EMA prompts. Participants were more than twice as likely to report engaging in sports or exercise before CS-EMA No-Activity prompts compared to R-EMA prompts. Participants were almost 5 times as likely to report going somewhere (ie, active or motorized transit) before CS-EMA Activity than R-EMA prompts but were half as likely to report going somewhere before CS-EMA No-Activity than R-EMA prompts. Participants were twice as likely to report sleeping before CS-EMA No-Data prompts than R-EMA prompts. As compared with answered R-EMA prompts, participants were more likely to report using technology before CS-EMA No-Activity prompts and less likely to report using technology before CS-EMA No-Data and CS-EMA Activity prompts. Participants were one and a half times as likely to report other activities (eg, getting ready, working part-time) before CS-EMA No-Activity than R-EMA. There were no significant BS effects for any activity type.

EMA Data Provided During Waist Accelerometer Nonwear Periods

Between the hours of 3 pm and 9 pm on weekdays, participants wore the waist accelerometer on an average of 4.09 (SD=2.16) hours per day (68.2% of the time). On 39 (10.3%) weekdays, waist-worn accelerometers were not worn at all (ie, less than 1

minute of total wear time). On weekend days between the hours of 7 am and 9 pm, participants wore the waist activity monitor on an average of 7.65 (SD=4.73) hours per day (54.6% of the time). On 23 (15.9%) weekend days, waist-worn accelerometers were not worn at all. Across all participants, there was a total of 885 answered EMA prompts (M=1.58, SD=2.30 per person per day) during periods when the waist accelerometer was not worn in the past 30 minutes. Of these prompts, 74.7% were EMA signal-contingent, 0.6% were CS-EMA Activity, 10.1% were CS-EMA No-Activity, and 14.5% were CS-EMA No-Data. Data on duration and type of self-reported activity (provided through answered R-EMA + CS-EMA) were summarized during waist accelerometer nonwear periods up to 60 minutes before each EMA prompt. EMA data provided an average of 31.9 (SD=54.10) minutes per person per day of activity data during waist accelerometer nonwear time (ie, that would have been missed if waist accelerometers alone were used). There was significantly more activity data per person per day provided by EMA during waist accelerometer nonwear periods on weekends (M=41.97, SE=6.08) than on weekdays (M=28.52, SE=5.32; z=3.06, *P*<.01). During waist accelerometer nonwear time periods, R-EMA + CS-EMA captured an additional 21 self-reported physical activity episodes (about 1 per child) that would have been missed using the waist-worn waist accelerometer alone. During waist accelerometer wear time periods, 20 bouts of activity were detected by CS-EMA Activity prompts. Seventeen bouts contained at least 10 minutes of light-intensity physical activity (ie, activity that does not meet MVPA thresholds), and only 3 had at least 10 minutes of MVPA. Of 58 bouts of MVPA detected by waist accelerometer that lasted longer than 15 minutes, 3 were detected by CS-EMA No-Activity, and 1 was detected by CS-EMA No-Data.

Contextual Characteristics of EMA-Reported Sports and Exercise

In total, sports and exercise were reported during 108 of 2,795 (3.7%) answered R-EMA and CS-EMA prompts. At each prompt, participants reported the duration of each activity within the window queried by the EMA. Sports and exercise activities had an average reported duration of 38.7 (SD=17.6 min) minutes before the EMA prompt. On average, the 30-minute windows before R-EMA and CS-EMA prompts with reported sports and exercise had 1.80 (SD=3.81) minutes of MVPA and 3.37 (SD=3.90) minutes of sedentary activity (measured by the waist-worn activity monitor). While performing that exercise or sport, children reported with the following frequencies that the mobile phone was being worn on their belt (2.8%), in their pocket (31.1%), in their handbag, purse, or backpack (21.7%), or that they were holding it in their hand (14.2%). Conversely, during sports and exercise, it was reported that the mobile phone was within reach but not being worn on the participant (9.4%) and not with the participants (20.8%). Reasons reported for not having the mobile phone (ie, *Not with me*) during sports and exercise were as follows: the battery died (4.5%), not wanting to damage it (59.1%), feeling it was too bulky (4.5%), feeling it was too uncomfortable (27.3%), and not being allowed to carry it (4.5%).

Table 5. Contextual characteristics of EMA-reported sports or exercise.

Characteristic	n (%)[b]
Social context	
Alone	9 (22.0)
Not alone	32 (78.0)
Physical context	
Outdoors	30 (69.8)
Indoors	13 (30.2)
Elevation change	
Mainly staying on flat ground	38 (76.0)
Mainly going uphill	4 (8.0)
Mainly going downhill	4 (8.0)
Going both uphill and downhill	4 (8.0)
Purpose	
For work, homework, or housework	1 (2.2)
Fun or recreation	28 (62.2)
Personal care	5 (11.1)
To get somewhere	2 (4.4)
Other	9 (20.0)
Baby sitting or childcare	0
Load bearing	
More than 20 lbs	3 (7.3)
10-20 lbs	1 (2.4)
5-10 lbs	4 (9.8)
Less than 5 lbs	3 (7.3)
None	20 (48.8)
Exercise form[a]	
None	3 (8.6)
Flexibility	19 (54.3)
Endurance	25 (71.4)
Balance	13 (37.1)
Strengthening	0
Type of exercise or sport	
Basketball, football, or soccer	16 (15.1)
Bicycling	3 (2.8)
Exercise, dance, or karate class	15 (14.2)
Other sports (baseball, skateboarding, and so forth)	27 (25.5)
Other running or jogging	20 (18.9)
Swimming	1 (0.9)
Walking	16 (15.1)
Weightlifting or strength training	8 (7.6)

[a]Participants could select more than one form of exercise.

[b]n=35 total responses.

Table 5 summarizes contextual and situational characteristics of sports and exercise reported through R-EMA + CS-EMA. The most common type of exercise or sport reported was *Other sports*, followed by *Other running/jogging, Basketball/Football/Soccer*, and *Walking*. Children most frequently reported that the purpose of performing the sport or exercise was for *Fun/Recreation*, and the most common form was endurance activities. Most children did not carry any extra weight during the sport or exercise, but a small percentage reported carrying over 20 lbs. Almost three-quarters of the activity occurred on flat ground (ie, not going uphill or downhill). Over two-thirds of the sports and exercise activities occurred outdoors, and a quarter occurred alone.

Discussion

Children recruited into surveillance, epidemiological, and intervention studies will increasingly have mobile phones, which are miniature computers with built-in motion sensors and electronic survey administration capabilities. Mobile phones are becoming more ubiquitous, affordable, and easy to use. The mobile phones are rarely far from the adolescents, and their affinity for the mobile phones creates new opportunities for real-time monitoring. This study tested the feasibility, performance, and utility of the Mobile Teen EMA app for capturing information about physical activity behaviors and contexts in adolescents.

Overall, participants answered over 80% of EMA surveys that were prompted, of which over 80% were answered after the first prompt. Relative to prior EMA work, this compliance could be considered moderate to high [27,28], especially because participants had the option to mute the mobile phone's audio. However, about a quarter of scheduled EMA surveys were never prompted because the mobile phone was either powered off or experienced an unknown technical problem. It is not clear whether participants intentionally turned off their mobile phones because they did not want to receive EMA prompts during those times, or whether the mobile phone battery died. Either way, EMA studies should make efforts to encourage participants to keep mobile phones powered on and fully charged during the hours of the days when EMA prompting is occurring, and EMA apps that use the mobile phone's sensing capabilities will need to be designed so that the sensing does not lead to additional battery drain and subsequent data loss. On average, EMA question sequences required less than 1 minute to complete, and participants completed surveys 20 seconds faster at the end of the 14-day monitoring period than at the beginning of the study—indicating improved ease and comfort with the survey items and procedures with practice. Furthermore, participants fully completed (ie, answered all questions) over 98% of the EMA surveys they started.

Adolescents may have been more likely to respond to CS-EMA than R-EMA prompts because CS-EMA prompts were triggered by a change in the state of the mobile phone device (eg, increase or decrease in movement), indicating that the participant was interacting with the mobile phone. In contrast, R-EMA may have been more likely to occur at times when the adolescent was away from the mobile phone (eg, sleep, bathing). The decreasing EMA prompt compliance rate across the 14-day monitoring period, however, suggests that some participant fatigue or growing disinterest may have occurred. Moreover, the lower EMA compliance rate in the mornings as compared with later times in the day may reflect participants' inability (due to sleep) or reluctance to answer EMA prompts in the early mornings on weekends. Girls may have been less likely to carry the mobile phones on their bodies than boys because the mobile phones did not fit as well in their pant pockets. Because girls are less likely to carry the mobile phone so that it has physical contact with the body, the built-in mobile phone accelerometers may be less likely to capture their fine-grained body movement patterns. However, the mobile phone may still have had the ability to detect major transitions in body movement patterns if the mobile phone was usually *"Within reach,"* as was reported (eg, sitting on a table nearby), which would suggest that girls tended to carry it with them when they make transition in location (even when indoors).

A second objective of this study was to evaluate the performance of the sensor-informed CS-EMA prompt triggers by examining differences in reported and objectively measured activity levels across the different EMA prompt types. The fact that levels of physical activity independently detected by the waist-worn accelerometer activity monitor were significantly higher during the time periods immediately leading up to CS-EMA Activity type prompts as compared with all other types of EMA prompts provides further support that full body movement (not just mobile phone movement) was indeed elevated during those periods. However, it is important to note that CS-EMA prompts did not capture all physical activity episodes that occurred. The CS-EMA Activity threshold was likely lower than the MVPA threshold applied to waist accelerometer data, as evidenced by most CS-EMA Activity prompts no more than light physical activity. Furthermore, participants were likely not in possession of their mobile phones during MVPA activities lasting longer than 15 minutes, and the CS-EMA No-Activity detection period was too long to capture this period of time because there were no bouts of MVPA longer than 60 minutes. Sports and exercise were also reported during CS-EMA No-Activity and CS-EMA No-Data type prompts. These findings are not surprising in light of the fact that during reported sports and exercise bouts, adolescents indicated that mobile phones were *"Within reach"* (but not being worn on the participant or not with the participants) over 30% of the time. Although CS-EMA Activity prompts may miss some sports and exercise owing to noncarrying of the mobile phone by adolescents, these bouts may be captured instead by the CS-EMA No-Activity and CS-EMA No-Data type prompts when the participants return to their mobile phones or turn them on again after physical activity. It is interesting to note that *"Going somewhere"* was 5 times as likely to occur before CS-EMA Activity prompts compared to R-EMA prompts, suggesting that CS-EMA Activity type prompting may be a good method to capture physically active travel among adolescents (eg, walking or bicycling for transportation).

The third and fourth objectives of this study address the utility of CS-EMA informed by onboard mobile phone motion sensors in providing information about physical activity and sedentary

behavior that would not otherwise be captured through a waist-worn accelerometer alone. One of the benefits of EMA is that it has the potential to yield activity data during periods when the waist-worn accelerometer is not being worn. Most adolescents are highly motivated to carry and keep charged and operate their own personal mobile phones. The internal motion sensor data from the mobile phones should therefore capture major transitions throughout the day, regardless of whether the waist-worn accelerometer is being worn. Even when an adolescent fails to have either device (waist-worn accelerometer or mobile phone) on his or her body, such as during high-contact sports and swimming, the CS-EMA feature should automatically trigger a CS-EMA No-Activity or CS-EMA No-Data survey prompt immediately after the adolescent begins carrying or using the mobile phone again. Thus, the CS-EMA may capture information about exercise or sports even if the mobile phone is not carried during these activities. Data about the type and duration of activities performed collected through CS-EMA can be used to estimate energy expenditure during accelerometer nonwear periods. Activity categories selected through CS-EMA reporting what the participant did during accelerometer nonwear periods can be converted to METs using the Compendium of Physical Activities [29] and multiplied by the duration of known device nonwear (in minutes) to generate an estimate of energy expenditure (in MET·minutes) for that period of time. These energy expenditure estimates can then be imputed to fill nonwear holes in objective activity data to obtain a more accurate representation of levels of physical activity and sedentary behavior across that day. In this study, EMA-reported activity type data were provided during about 32 minutes per day of waist-worn accelerometer nonwear.

Additional information about physical activity episodes provided by EMA that would not otherwise be captured through a waist-worn accelerometer includes situational and contextual characteristics. Data from the sensor-informed EMA can be used to improve energy expenditure estimates for activities not well captured by waist-worn motion sensors, such as those that involve the upper body (eg, weightlifting or strength training was reported in almost 8% of sports or exercise bouts), cycling (reported in almost 3% of sports or exercise bouts), weight bearing (some weight reported being carried on over 50% of sports or exercise bouts), and incline or decline (reported in almost 25% of sports or exercise bouts). These data can be used to upwardly or downwardly adjust energy expenditure estimates obtained from objective activity monitors [11,30]. Furthermore, EMA data may also be used to differentiate between conceptually distinct activity types (eg, exercise, dance, or karate class [14% of sports or exercise bouts] vs running or jogging [19% of sports or exercise bouts]), which may appear identical when examining objective activity intensity data alone. In addition, EMA gathers data about the purpose of the activity (eg, fun or recreation [62% of sports or exercise bouts], to get somewhere [4% of sports or exercise bouts], for work or housework [2% of sports or exercise bouts]) that may be useful in assessing the amount of transit- and work-related physical activity performed. Finally, the EMA questions gather information about where, with whom, and why physical activity occurs, as well as how participants feel during those activities. These data help researchers to understand whether physical

activity intensity or duration differs across contexts [31] and to investigate time-varying antecedents and consequences of behavior [27].

One possible concern with the method as proposed is that the Mobile Teen app depends on adolescents in future activity measurement studies using personal mobile phones. The mobile phones they have may not be appropriate mobile phones for running the Mobile Teen app. In those cases, some of the adolescents could be switched to appropriate mobile phones by temporarily swapping SIM cards, as was done in the Mobile Teen testing (for approximately 20% of participants). The technology in its current form will only work on Android mobile phones because iOS will not support the required background processing, but over 80% of new mobile phone shipments worldwide use Android [32], and changes to Apple's iPhone line adding a motion coprocessor chip may allow continuous movement measurement [33] and thereby create opportunities to develop versions of Mobile Teen for new iPhones as well. CS-EMA activity and duration thresholds could also be modified in future iterations to better capture episodes of physical activity for a given population, especially when the participant is not in possession of a mobile phone. It should be noted that the assessment time frame used in this study (3-9 pm on weekdays and 7 am-9 pm on weekend days) did not permit an overall level of physical activity to be measured (because physical activity taking place on the way to school or during physical education classes was not included). Another concern often raised with EMA is reactivity, the potential for behavior to be impacted by the very act of assessing it, but the magnitude of reaction to EMA has been observed to be small for EMA studies [34-36]. Another limitation is the ~20% of prompted EMA surveys that are unanswered, which are more common in the mornings. Reasons for lower compliance rates in non-Hispanic (vs Hispanic) children are largely unknown and should be explored in future research. However, it should be noted that Hispanic children comprised the majority (66%) of the current sample. There may be other unmeasured variables that correlate with noncompliance to EMA prompting such as negative mood and stress, which should be explored in future studies. Furthermore, this study tested the Mobile Teen app on a relatively small sample of primarily Hispanic adolescents from an urban Los Angeles high school. Further testing is needed in larger samples of adolescents from other regions of the United States and internationally.

Ultimately, the sensor-informed CS-EMA methods used by the Mobile Teen app can be used to augment and supplement physical activity data collected through external objective activity monitors (such as waist or wrist-worn accelerometers). Studies may deploy CS-EMA procedures in a stand-alone manner or in conjunction with objective activity monitoring, depending on which characteristics of physical activity are desired for assessment. For example, if a study seeks to understand the types (and contexts) of physical activity that result in bouts of high-intensity activity that last 15+ minutes, sensor-informed CS-EMA procedures could be used alone without an external objective activity monitor. In addition, sensor-informed CS-EMA may be used together with an external objective activity monitor to capture and estimate energy

XSL•FO

RenderX

expenditure during times when the external objective activity monitor is not worn. Furthermore, the sensor-informed CS-EMA procedures described in this study can provide the basic architecture for Just-In-Time Adaptive Interventions (JITAIs) [12,37,38], targeting physical activity change. JITAIs use real-time decision rules to link a participant's current situation, sometimes inferred from sensor data, with appropriate, tailored intervention strategies intended to have optimal impact. Sensor-informed CS-EMA procedures can be built into the learning phases of JITAIs to collect information about affective, motivational, and contextual factors that are ideographically related to naturally occurring physical activity variations and inform the development of intervention content and messages. Sensor-informed CS-EMA can also guide the timing during the subsequent intervention delivery phase of JITAIs by prompting individuals to break up elongated bouts of sedentary behavior or to lengthen short bouts of higher intensity physical activity detected by the mobile phone. The opportunity to move beyond randomly delivered behavior change messages to an informed and timely delivery strategy has enormous potential benefits in the intervention domain.

Acknowledgments

Conflicts of Interest

None declared.

References

1. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008 Jan;40(1):181-188. [doi: 10.1249/mss.0b013e31815a51b3] [Medline: 18091006]

2. Metcalf B, Henley W, Wilkin T. Effectiveness of intervention on physical activity of children: Systematic reviewmeta-analysis of controlled trials with objectively measured outcomes (EarlyBird 54). BMJ 2012;345:e5888.

3. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U, Lancet Physical Activity Series Working Group. Global physical activity levels: surveillance progress, pitfalls, and prospects. Lancet 2012 Jul 21;380(9838):247-257. [doi: 10.1016/S0140-6736(12)60646-1] [Medline: 22818937]

4. Haskell WL. Physical activity by self-report: a brief history and future issues. J Phys Act Health 2012 Jan;9 Suppl 1:S5-10. [Medline: 22287448]

5. Ainsworth BE, Caspersen CJ, Matthews CE, Mâsse LC, Baranowski T, Zhu W. Recommendations to improve the accuracy of estimates of physical activity derived from self report. J Phys Act Health 2012 Jan;9 Suppl 1:S76-S84 [FREE Full text] [Medline: 22287451]

6. Hallal P, Reichert F, Clark V, Cordeira K, Menezes A, Eaton S. Energy expenditure compared to physical activity measured by accelerometry and self-report in adolescents: A validation study. PLoS One 2013;8(11):e77036.

7. Wells SL, Kipping RR, Jago R, Brown J, Hucker D, Blackett A, et al. Characteristics associated with requested and required accelerometer wear in children. BMJ Open 2013;3(8):e003402 [FREE Full text] [doi: 10.1136/bmjopen-2013-003402] [Medline: 23975106]

8. Tudor-Locke C, Mire EF, Dentro KN, Barreira TV, Schuna JM, Zhao P, et al. A model for presenting accelerometer paradata in large studies: ISCOLE. Int J Behav Nutr Phys Act 2015;12:52 [FREE Full text] [doi: 10.1186/s12966-015-0213-5] [Medline: 25927615]

9. Ottevaere C, Huybrechts I, De MF, De BI, Cuenca-Garcia M, De HS. The use of accelerometry in adolescents and its implementation with non-wear time activity diaries in free-living conditions. J Sports Sci 2011 Jan;29(1):103-113. [doi: 10.1080/02640414.2010.521169] [Medline: 21104522]

10. Freedson P, Bowles HR, Troiano R, Haskell W. Assessment of physical activity using wearable monitors: recommendations for monitor calibration and use in the field. Med Sci Sports Exerc 2012 Jan;44(1 Suppl 1):S1-S4 [FREE Full text] [doi: 10.1249/MSS.0b013e3182399b7e] [Medline: 22157769]

11. Butte NF, Ekelund U, Westerterp KR. Assessing physical activity using wearable monitors: measures of physical activity. Med Sci Sports Exerc 2012 Jan;44(1 Suppl 1):S5-12. [doi: 10.1249/MSS.0b013e3182399c0e] [Medline: 22157774]

12. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? Transl Behav Med 2011 Mar;1(1):53-71 [FREE Full text] [doi: 10.1007/s13142-011-0021-7] [Medline: 21796270]

13. Nahum-Shani S, Smith SN, Tewari A, Witkiewitz K, Collins LM, Spring B, et al. Just in time adaptive interventions (JITAIs): An organizing framework for ongoing health behavior support. In: The Methodology Center (Technical Report No. 14-126). Penn State: The Methodology Center; 2014.

14. Shiffman S, Stone A, Hufford M. Ecological Momentary Assessment. Annu. Rev. Clin. Psychol 2008 Apr;4(1):1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415]

15. Kaplan RM, Stone AA. Bringing the laboratory and clinic to the community: mobile technologies for health promotion and disease prevention. Annu Rev Psychol 2013;64:471-498. [doi: 10.1146/annurev-psych-113011-143736] [Medline: 22994919]

16. Shiffman S. Designing protocols for Ecological Momentary Assessment. New York: Oxford Press; 2007.

17. Schüz N, Cianchi J, Shiffman S, Ferguson S. Novel technologies to study smoking behavior: Current developments in ecological momentary assessment. Current Addiction Reports 2015;2(1):8-14.

18. Schwartz N. RetrospectiveConcurrent Self-Reports: The Rationale for Real-Time Data Capture. In: Stone AS, Shiffman SS, Atienza A, Nebeling L. editors. The Science of Real-Time Data Capture: Self-Report in Health Research. Oxford, UK: Oxford University Press,; 2007:76-113.

19. Intille S. Technological innovations enabling automatic, context-sensitive ecological momentary assessment. In: Stone AA, Shiffman, S., Atienza, A. A., Nebeling, L. editor. The Science of Real-Time Data Capture: Self-Report in Health Research. New York: Oxford University Press; 2007:308-337.

20. Bort-Roig J, Gilson ND, Puig-Ribera A, Contreras RS, Trost SG. Measuring and influencing physical activity with smartphone technology: a systematic review. Sports Med 2014 May;44(5):671-686. [doi: 10.1007/s40279-014-0142-5] [Medline: 24497157]

21. Donaire-Gonzalez D, de NA, Seto E, Mendez M, Nieuwenhuijsen MJ, Jerrett M. Comparison of physical activity measures using mobile phone-based CalFit and Actigraph. J Med Internet Res 2013;15(6):e111 [FREE Full text] [doi: 10.2196/jmir.2470] [Medline: 23896156]

22. Dunton GF, Dzubur E, Kawabata K, Yanez B, Bo B, Intille S. Development of a smartphone application to measure physical activity using sensor-assisted self-report. Front Public Health 2014;2:12 [FREE Full text] [doi: 10.3389/fpubh.2014.00012] [Medline: 24616888]

23. Freedson P, Pober D, Janz KF. Calibration of accelerometer output for children. Med Sci Sports Exerc 2005 Nov;37(11 Suppl):S523-S530. [Medline: 16294115]

24. Bryk A, Raudenbush S. Hierarchical linear models: applications and data analysis methods. Newbury Park, CA: Sage; 1992.

25. Curran PJ, Bauer DJ. The disaggregation of within-person and between-person effects in longitudinal models of change. Annu Rev Psychol 2011;62:583-619 [FREE Full text] [doi: 10.1146/annurev.psych.093008.100356] [Medline: 19575624]

26. Dzubur E. EconPapers. 2015. XTCENTER: Stata module to disaggregate within and between-person effects by centering variables for mixed and melogit models URL: http://econpapers.repec.org/software/bocbocode/s458076.htm [accessed 2016-04-27] [WebCite Cache ID 6h5U5ohZr]

27. Dunton GF, Huh J, Leventhal AM, Riggs N, Hedeker D, Spruijt-Metz D, et al. Momentary assessment of affect, physical feeling states, and physical activity in children. Health Psychol 2014 Mar;33(3):255-263 [FREE Full text] [doi: 10.1037/a0032640] [Medline: 23668846]

28. Litcher-Kelly L, Stone AA, Broderick JE, Schwartz JE. Associations among pain intensity, sensory characteristics, affective qualities, and activity limitations in patients with chronic pain: a momentary, within-person perspective. J Pain 2004 Oct;5(8):433-439. [doi: 10.1016/j.jpain.2004.07.005] [Medline: 15501425]

29. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Tudor-Locke C, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. Med Sci Sports Exerc 2011 Aug;43(8):1575-1581. [doi: 10.1249/MSS.0b013e31821ece12] [Medline: 21681120]

30. Uyanik I, Khatri A, Majeti D, Ugur M, Shastri D, Pavlidis I, editors. Using accelerometer data to estimate surface inclineits walking app potential. 2015 Presented at: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems; : ACM; 2015; Seoul, Republic of Korea.

31. Dunton GF, Liao Y, Intille S, Wolch J, Pentz MA. Physical and social contextual influences on children's leisure-time physical activity: an ecological momentary assessment study. J Phys Act Health 2011 Jan;8 Suppl 1:S103-S108. [Medline: 21350250]

32. IDC. IDC Research. 2015. Smartphone OS Market Share, 2015 Q2 URL: http://www.idc.com/prodserv/smartphone-os-market-share.jsp [accessed 2016-04-27] [WebCite Cache ID 6h5UMXXcY]

33. Georgiev P. Energy Efficientfair management of sensing applications on heterogeneous resource mobile devices. 2015 Presented at: Proceedings of the 2015 on MobiSys PhD Forum; 2015; Florence, Italy.

34. Cruise CE, Broderick J, Porter L, Kaell A, Stone AA. Reactive effects of diary self-assessment in chronic pain patients. Pain 1996 Oct;67(2-3):253-258. [Medline: 8951918]

35. Hufford MR, Shields AL, Shiffman S, Paty J, Balabanis M. Reactivity to ecological momentary assessment: an example using undergraduate problem drinkers. Psychol Addict Behav 2002 Sep;16(3):205-211. [Medline: 12236455]

36. McCarthy DE, Minami H, Yeh VM, Bold KW. An experimental investigation of reactivity to ecological momentary assessment frequency among adults trying to quit smoking. Addiction 2015 Oct;110(10):1549-1560. [doi: 10.1111/add.12996] [Medline: 26011583]

37. Sharmin M, Raij A, Epstien D, Nahum-Shani I, Beck J, Vhaduri S. Visualization of time-series sensor data to inform the design of just-in-time adaptive stress interventions. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2015 Presented at: ACM; 2015; Osaka, Japan.

38.   Kumar S, Nilsen W, Pavel M, Srivastava M. Mobile Health: Revolutionizing Healthcare Through Transdisciplinary Research. Computer 2013 Jan;46(1):28-35. [doi: 10.1109/MC.2012.392]

Abbreviations

BMI: body mass index
BS: between subjects
EMA: ecological momentary assessment
CI: confidence interval
CS-EMA: context-sensitive ecological momentary assessment
GSM: global system for mobile communication
JITAI: Just-In-Time Adaptive Intervention
MET: metabolic equivalent
MVPA: moderate-to-vigorous physical activity
OR: odds ratio
R-EMA: random ecological momentary assessment
SD: standard deviation
SE: standard error of the mean
SIM: subscriber identity module
SMS: short message service
WS: within subjects

XSL•FO
RenderX

<u>Original Paper</u>

Completeness and Reliability of Location Data Collected on the Web: Assessing the Quality of Self-Reported Locations in an Internet Sample of Men Who Have Sex With Men

Adam S Vaughan[1,2], M.P.H., M.S.; Michael R Kramer[1,2], Ph.D.; Hannah LF Cooper[2,3], ScD; Eli S Rosenberg[1,2], Ph.D.; Patrick S Sullivan[1,2], DVM, Ph.D.

[1]Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, United States

[2]Laney Graduate School, Emory University, Atlanta, GA, United States

[3]Department of Behavioral Science and Health Education, Rollins School of Public Health, Emory University, Atlanta, GA, United States

Corresponding Author:
Adam S Vaughan, M.P.H., M.S.
Department of Epidemiology
Rollins School of Public Health
Emory University
1518 Clifton Road NE
Atlanta, GA, 30322
United States
Phone: 1 404 727 2038
Fax: 1 404 727 8661
Email: adam.s.vaughan@gmail.com

Abstract

Background: Place is critical to our understanding of human immunodeficiency virus (HIV) infections among men who have sex with men (MSM) in the United States. However, within the scientific literature, place is almost always represented by residential location, suggesting a fundamental assumption of equivalency between neighborhood of residence, place of risk, and place of prevention. However, the locations of behaviors among MSM show significant spatial variation, and theory has posited the importance of nonresidential contextual exposures. This focus on residential locations has been at least partially necessitated by the difficulties in collecting detailed geolocated data required to explore nonresidential locations.

Objective: Using a Web-based map tool to collect locations, which may be relevant to the daily lives and health behaviors of MSM, this study examines the completeness and reliability of the collected data.

Methods: MSM were recruited on the Web and completed a Web-based survey. Within this survey, men used a map tool embedded within a question to indicate their homes and multiple nonresidential locations, including those representing work, sex, socialization, physician, and others. We assessed data quality by examining data completeness and reliability. We used logistic regression to identify demographic, contextual, and location-specific predictors of answering all eligible map questions and answering specific map questions. We assessed data reliability by comparing selected locations with other participant-reported data.

Results: Of 247 men completing the survey, 167 (67.6%) answered the entire set of eligible map questions. Most participants (>80%) answered specific map questions, with sex locations being the least reported (80.6%). Participants with no college education were less likely than those with a college education to answer all map questions (prevalence ratio, 0.4; 95% CI, 0.2-0.8). Participants who reported sex at their partner's home were less likely to indicate the location of that sex (prevalence ratio, 0.8; 95% CI, 0.7-1.0). Overall, 83% of participants placed their home's location within the boundaries of their reported residential ZIP code. Of locations having a specific text description, the median distance between the participant-selected location and the location determined using the specific text description was 0.29 miles (25th and 75th percentiles, 0.06-0.88).

Conclusions: Using this Web-based map tool, this Web-based sample of MSM was generally willing and able to provide accurate data regarding both home and nonresidential locations. This tool provides a mechanism to collect data that can be used in more nuanced studies of place and sexual risk and preventive behaviors of MSM.

(J Med Internet Res 2016;18(6):e142)   doi:10.2196/jmir.5701

XSL•FO
RenderX

KEYWORDS

HIV; digital mapping; geographic locations; survey; men who have sex with men

Introduction

Place, or the context simultaneously experienced and defined by individuals [1], is critical to our understanding of human immunodeficiency virus (HIV) among men who have sex with men (MSM) in the United States. Through surveillance data, place fundamentally shapes our understanding of the epidemiology of the epidemic [2]. As a contextual exposure, place represents both a foundational environment in which HIV-related behaviors occur and a potential modifier of the pathway between other contextual exposures and HIV-related outcomes [3,4]. However, within the public health literature, place is almost always defined as a residential location [5-13], suggesting a fundamental assumption of equivalency between place of residence, place of sexual risk, and place of prevention. US national HIV case surveillance data make the same assumption, most often reporting data based on residence at the time of diagnosis [2].

Despite this implicit assumption, HIV-related sexual risk and prevention behaviors of MSM do not necessarily occur within the residential neighborhood [14-19]. Social ecologic theory acknowledges the importance of nonresidential locations (such as the broader urban environment and gay venues) in determining these behaviors [3,4,20-23]. For example, the availability of HIV testing services and venues where MSM gather may be influenced by broader social characteristics and norms. Access to these services and venues may then influence the formation of sexual networks and promote or inhibit individual-level behaviors, such as regular HIV testing and unprotected sex [3,4,23,24]. Consequently, using only residential neighborhood as a proxy for the many levels of sociocontextual factors may miss critical health-related exposures. To address this potential misclassification, the concept of activity spaces, which represent the collection of locations to which an individual has been exposed, has recently been introduced into the HIV literature [14,15,25].

Measuring activity spaces requires collecting large amounts of detailed geographic data. Prior studies have used global positioning systems (GPS) [26-31] or interviewer-assisted means to establish specific locations and, ultimately, to measure activity spaces [14,15]. Although these methods provide a precise and comprehensive set of locations, they have limitations. Collecting locations with GPS requires processing large amounts of data and a large investment in purchasing and maintaining the GPS devices. Interviewer-assisted methods require a large time and budget commitment, limiting the number of potential participants in a study.

To begin to address these limitations, our research group recently developed a Web-based tool that allows participants to select locations using a Google Maps question embedded within a Web-based survey [32]. Given the potentially sensitive nature of these data, participants may be more comfortable reporting such data in an anonymous Web-based survey [33]. In validation of this Web-based tool using home and health care provider locations among a cohort of HIV-positive Atlanta-area MSM, approximately 84% of participants indicated these locations using the map-based tool [32]. Among participants recruited on the Web, 50% of locations entered using the map-based tool were found to be within 0.3 miles of the true location (interquartile range, 0.1-1.1 miles). However, this previous study collected data for a limited number of locations from a population defined by a single geographic area (Atlanta, Georgia) and health status (HIV positive). Because research participation may differ by demographic and health-related factors, these results may not be generalizable to a broader population of MSM [34-37].

Therefore, given the need to gather detailed spatial data for HIV-related behaviors among MSM, to overcome current challenges in its collection, and to expand on prior validation efforts, this study examines the quality of spatial data collected using a Web-based map tool. Specifically, using a Web-based map tool to collect both residential and relevant nonresidential locations (eg, sex locations, HIV testing, work, socialization), this study examines the completeness and reliability of data collected from MSM living in a wide range of geographic locations and independent of HIV status.

Methods

Recruitment

Participants were recruited using Facebook banner ads, a method that has been shown to yield samples with similar risk behaviors and demographics (excepting race) as venue-based methods of recruiting MSM [37]. Ads were targeted to users based on geography and interests. A $3 donation to a charity the participant selected from a predefined list was provided as incentive.

Eligible participants were required to be male at birth, aged 18 years or older, be able to read and write English, and had to report at least one male sex partner in the past 6 months and to reside in Georgia, Texas, or Wisconsin. These 3 states vary in their underlying HIV epidemiology, demographics, and contextual factors, which could be associated with willingness to answer our map questions and allowed us to draw conclusions based on a diverse convenience sample of MSM. This population also expands on the population used in the prior validation of this tool [32]. Participants who met eligibility criteria completed a Web-based consent form.

Collection of Place-Based Data

Consenting participants completed a Web-based survey that included demographic and behavioral questions and an item on residential ZIP code at the time of data collection.

In addition to these questions, participants were asked to use a map-based tool (Figure 1) [32] to drop a pin onto a Google map to indicate the following specific locations that may be relevant to the daily lives and health-related behaviors of MSM: home; work or school location, if the participant reported working at

XSL•FO

RenderX

least part time or being a student; locations of up to 3 sexual encounters in the past 6 months; locations of up to 2 socialization locations; location of last HIV test, within the past year; location of the last test for another sexually transmitted infection, within the past year; primary care physician, if the participant reported having a primary care physician; pharmacy, if the participant reported having a regular pharmacy; and location where he received free condoms, if the participant reported picking up free condoms in the past 6 months.

For each location of interest, participants could choose to not answer the map question and were asked to indicate why they chose not to answer. These reasons were then categorized as either unable or unwilling to answer the question. Answer options indicating that a participant was unable to select the location were the following: "I can't remember where this location is," "I'm not sure where that place is on a map," "I'm not comfortable using the map to select locations," "This place is in a different city." Answer options indicating that a participant was unwilling to select the location were "Didn't feel comfortable giving that information," "Worried about a loss of privacy," "Worried about what friends, family, or coworkers would think."

Participants were also allowed to indicate that a location was the same as another previously reported location (eg, report sex at home). In these cases, participants were not required to select the location a second time or to indicate a reason for not selecting the location. Willingness to use the map-based tool to answer the second location was assumed the same as that of the previously reported location.

For many types of locations, participants needed to report engaging in a qualifying behavior to be eligible to answer the corresponding map-based question. For example, participants needed to report having a regular physician before being presented with the map to identify physician location. As a result, the number of participants eligible to answer each location question varied.

In addition, for each location, participants entered a name that was used to reference that location throughout the survey. This name was entered by participants and could be generic (eg, home, work, bar) or specific (eg, Dr. Smith, Walgreens).

Figure 1. Sample of Google Maps question embedded within the Web-based survey.
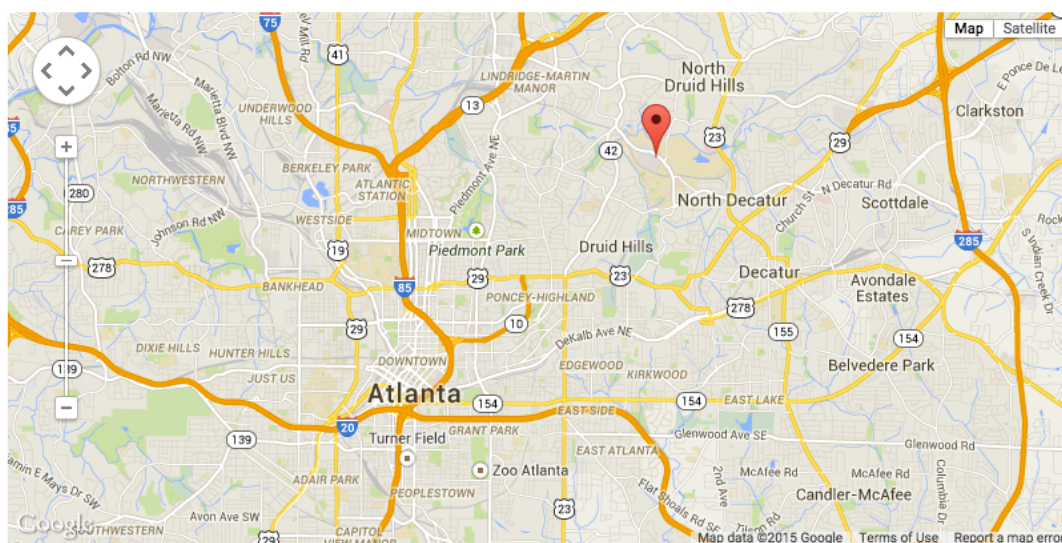
Primary Outcome Definitions

This analysis uses 2 different primary outcomes: answering the entire set of map questions and answering specific map questions. A participant was considered to have answered the entire set of map questions if he used the map-based tool to indicate all locations for which he was eligible to answer. More granularly, the second outcome required participants to indicate specific eligible locations (eg, home, socialization, sex) using the map-based tool.

Covariate Definitions

The covariates of interest in this study represent demographic variables, contextual factors related to residential location, and factors specific to given location types. All these factors could potentially be associated with an individual being unwilling or unable to answer the location-based questions.

Age was categorized into 3 groups with breaks at ages 25 and 50 years, in accordance with age group definitions used in the Centers for Disease Control and Prevention reporting of HIV surveillance data [38,39]. Due to a limited number of nonwhite participants, self-reported race was categorized as white or nonwhite. Education was categorized as high school diploma

or less, any college, or college degree. HIV status was self-reported. State was defined as the state where the participant reported currently living. Each participant was asked to indicate his primary mode of transportation, and this was dichotomized into primarily using a car and primarily using other, noncar transportation.

Residential poverty and residential urbanicity were defined based on the reported residential ZIP code. Poverty was defined using ZIP code tabulation areas (ZCTA) from the US Census Bureau's 2009-2013 five-year American Community Survey estimates and categorized as low poverty (<20% poverty), high poverty (≥20% poverty), or concentrated poverty (≥40% poverty), based on federal poverty definitions [40]. Urbanicity was defined using the 2013 National Center for Health Statistics Urban-Rural Classification Scheme for Counties [41], with the 2 most rural categories combined. For each sex location, participants reported the type of location (eg, sex partner's home). Participants also reported any condomless anal intercourse (CAI) at last sex at each reported sex location.

Statistical Analysis

Overview

After calculation of descriptive statistics for the covariates of interest, this analysis had 3 parts. We first examined factors associated with answering the entire set of map questions. Second, in an item-specific analysis, we examined factors associated with answering specific map questions (eg, home, sex locations). Finally, we examined the reliability of the reported locations.

Response to the Entire Set of Map Questions

Data regarding answering the entire set of map questions for which participants were eligible were first summarized by the covariates of interest. In bivariate analyses, we compared completeness across the levels of each covariate using chi-square and Fisher exact tests.

We then performed multivariable analyses to examine associations between the given covariates and answering all eligible map questions. Predictive margins methods were used with logistic regression to estimate adjusted prevalence ratios (PRs) for answering all map questions [42,43]. This method permitted direct estimation of adjusted PR, rather than an estimated prevalence odds ratio. Because we expected most men to respond to these questions (ie, the outcome is not rare) [32], the prevalence odds ratio estimated using logistic regression would be larger than the true PR and, consequently, direct estimation of the PR is preferred [44]. This method also avoids statistical convergence issues that may occur when estimating PR using other methods, such as log-binomial regression [45]. This model included the following possible predictors: age, race, poverty category for the residential ZIP code, residential urbanicity, state, education, HIV status, HIV test within the past year, and primary mode of transportation.

Response to Specific Map Questions

Data regarding answering specific map questions (ie, locations of home, sex, socialization, workor school, last HIV test, last sexually transmitted infection test, primary care physician, pharmacy, and free condoms) were first summarized by the covariates of interest. In bivariate analyses, we compared completeness in answering each type of map questions across the levels of each covariate using chi-square and Fisher's exact tests. Proportions of the reasons for nonresponse were calculated.

We again used predictive margins methods with logistic regression to examine associations between the covariates of interest and answering specific map questions. Nine models were created, one for each location type. Each model included the following possible predictors of prevalence of response: age, race, residential poverty, residential urbanicity, state, education, HIV status, HIV test within the past year, and primary mode of transportation. The model for sex locations also included CAI and sex at the partner's home. The model for reporting an HIV test location was restricted to HIV-negative participants. Each participant entered up to 2 socialization locations and up to 3 sex locations. Consequently, models for these 2 types of locations accounted for within-participant correlation using marginal models with exchangeable correlation structure.

Data Reliability

Data reliability was assessed using 2 methods. First, agreement between a reported ZIP code and residential location was determined. Other address information was not collected in this study. To measure this agreement, each residential location identified using the map tool was geocoded to a ZCTA, the US Census Bureau's representation of ZIP codes. Agreement between the geocoded ZCTA and the participant self-reported ZIP code was then defined as an exact match between the 2 values.

In addition, reliability was assessed using distances between the reported location and name of the reported location. In this study, we asked men to identify locations for which they may not readily know the addresses and, consequently, for which a formal validation was not possible within this study. Therefore, for each location, participants entered text to help them identify the location in additional questions about that location. Using this text and the type of location, a Google Maps search was completed around the location selected using the map tool. If this search was informative, the distance between the reported point and the actual point were recorded. If the participant-entered text was generic (eg, doctor), rather than a specific name (eg, Dr Smith), then the driving distance between the selected location and the nearest location matching that description was recorded. Distances were summarized by those matched by a generic name, those matched by a specific name, and those matched using only a geographic location.

Analysis Software

Data management was performed using SAS, v9.4 (SAS Institute, Cary, NC, USA). Geocoding and spatial data manipulation were completed in R, v3.2.1 (R Foundation for Statistical Computing, Vienna, Austria) [46]. Predictive margins models were performed using SAS-callable SUDAAN, v11.0.1 (Research Triangle Institute, Research Triangle Park, NC, USA).

Ethics

This study was approved by the Emory University Institutional Review Board (protocol #IRB00074519).

Results

Sample Characteristics and Question Completeness

Of 105,815 men presented with the Facebook ad, 3058 men (2.9%) clicked on the ad to enter the eligibility screening. Of these, 624 men (20.4%) were eligible, of whom 341 men (11.1% of those screened, 54.6% of those eligible) consented to participate in the study. 247 men (72.4%) completed the survey and are included in this analysis. Our sample represented a wide range of ages, urbanicity, and poverty levels (Table 1). Our sample was highly educated and largely white.

Table 1. Sample characteristics (N=247).

Covariate		Number (%)
Age (years)		
	18-25	66 (26.7)
	26-50	103 (41.7)
	51 and older	78 (31.6)
Race		
	White	202 (81.8)
	Non-white race	45 (18.2)
Reported HIV[a] positive		36 (14.6)
HIV[a] test within the past year[b]		119 (56.4)
Education		
	High school or less	22 (8.9)
	Some college	89 (36.0)
	College degree	136 (55.1)
State		
	Georgia	76 (30.8)
	Texas	134 (54.3)
	Wisconsin	37 (15.0)
Primary mode of transportation		
	Car	227 (91.9)
	Other	20 (8.1)
Residential poverty		
	Low	157 (63.6)
	High	71 (28.7)
	Concentrated	19 (7.7)
Urbanicity		
	Urban core	108 (43.7)
	Suburban	48 (19.4)
	Medium metro	41 (16.6)
	Small metro	31 (12.6)
	Nonmetropolitan	19 (7.7)

[a]HIV: human immunodeficiency virus.

[b]Among participants who do not report being HIV positive.

Response to the Entire Set of Map Questions

Of included participants, 167 (67.6%) answered all map questions for which they were eligible. Nine participants (3.6%) answered none of the map questions for which they were eligible. Of the remaining participants, 71 (28.7%) answered at least one, but not all, map questions.

In unadjusted analyses (Figure 2), only less education was associated with significantly less completion of all map questions (P<.001), with 31.8% of participants with a high

school diploma or less answering all questions, compared with 70.0% of participants with some college and 72.1% of participants with a college degree. This finding was confirmed in adjusted analyses (Figure 3), with participants with no college education being roughly half as likely as those with a college education to answer all eligible map questions (PR, 0.4; 95% CI, 0.2-0.8). No other covariate was significantly associated with answering all eligible map questions in unadjusted or adjusted analyses.

Figure 2. Crude percent of participants answering all eligible map questions and specific map questions. Percentages are the proportion of individuals within the given covariate level eligible to answer the map question who completed the given map question. Statistically significant differences are indicated in black filled circles.
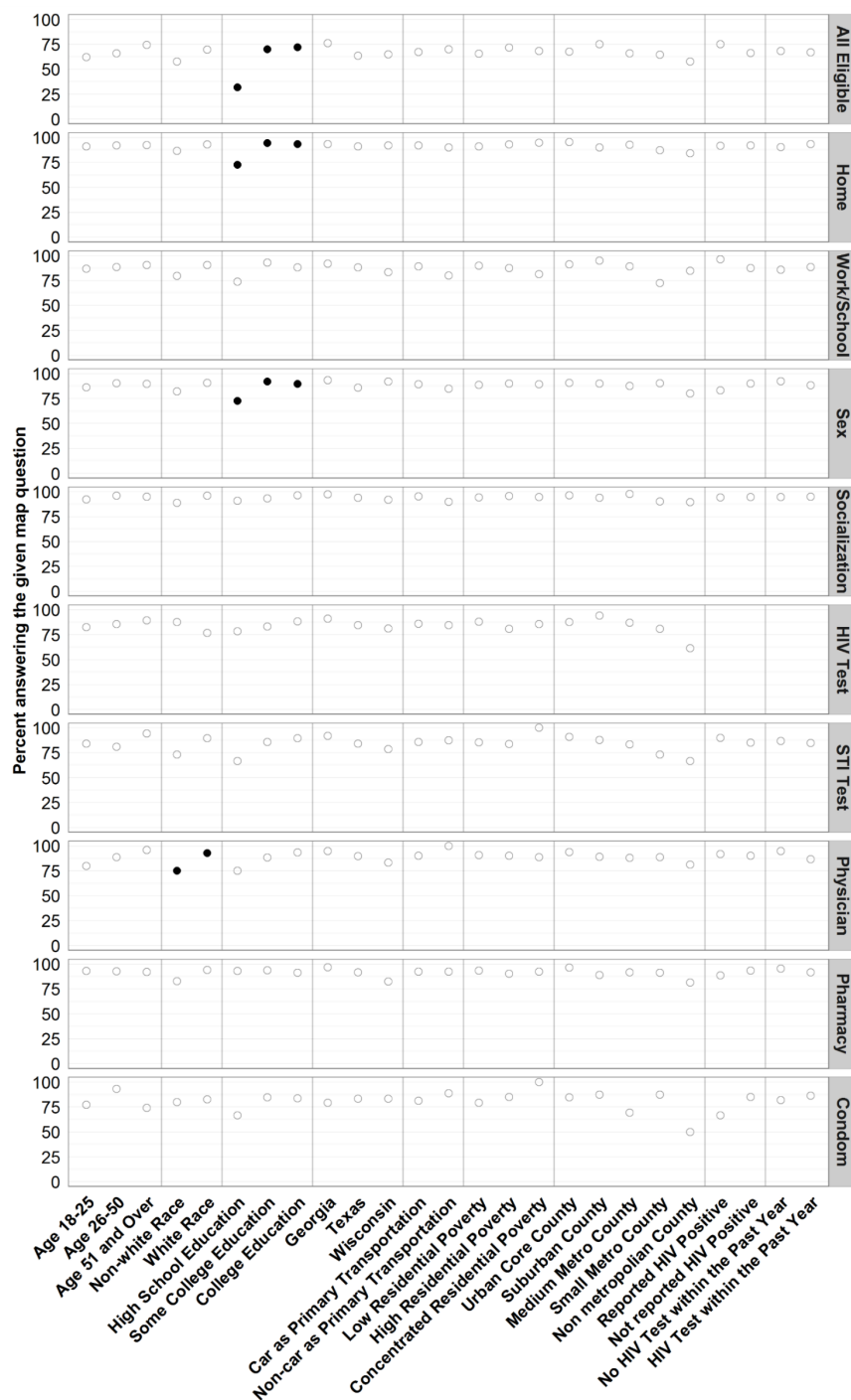
Figure 3. Adjusted prevalence ratios (aPR) and 95% CI for answering all eligible map questions and specific map questions by each covariate. aPRs are adjusted by all other covariates. Statistically significant aPRs are indicated with black filled circles. The scale of the y axis is logarithmic and differs across location types to better visualize the CI.
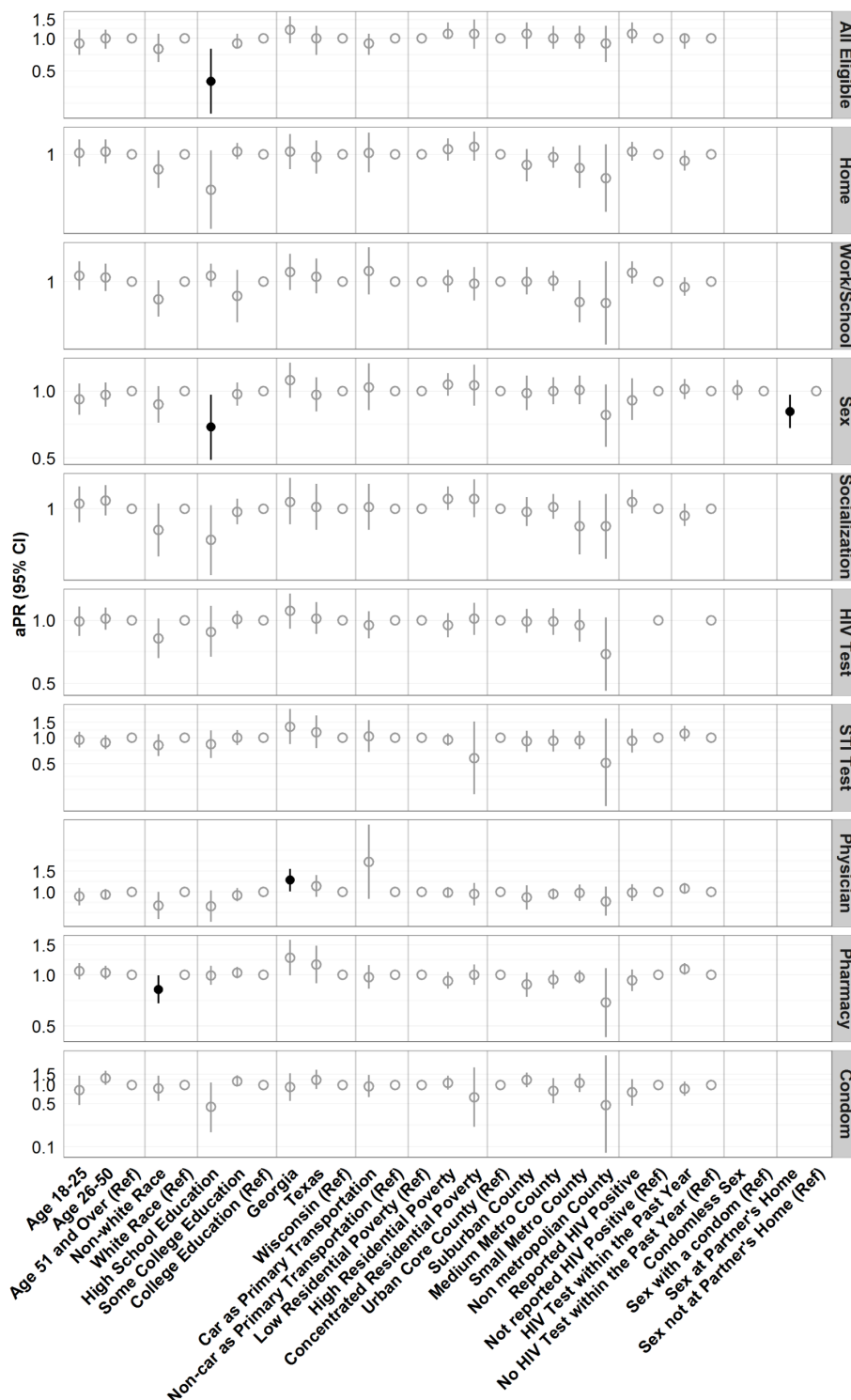
Response to Specific Map Questions

In item-specific analyses, most (>80%) of those eligible answered each individual map question (Table 2). Sex locations were the least likely to be answered (80.6%). For most locations, participants who chose to not answer the map-based question were generally unwilling to answer, rather than unable to answer (Table 2). However, for sex locations and HIV testing locations, the proportion of participants who were unable to answer was similar to the proportion who were unwilling to answer.

Table 2. Ability and willingness to answer specific map-based questions.

Location	Total eligible (%)	Answered (%)	Unable (%)	Unwilling (%)	Both unwilling and unable (%)	No reason given (%)
Home	247 (100)	227 (91.9)	2 (0.8)	15 (6.1)	3 (1.2)	0 (0)
Work or school	209 (84.6)	185 (88.5)	2 (1.0)	21 (10.0)	0 (0)	1 (0.5)
Socialization[a]	474 (96.0)	430 (90.7)	6 (1.3)	33 (7.0)	3 (0.6)	2 (0.4)
Sex[b]	396 (53.4)	319 (80.6)	30 (7.6)	36 (9.1)	3 (0.8)	8 (2.0)
HIV[c] test[d]	119 (56.4)	103 (86.6)	9 (7.6)	7 (5.9)	0 (0)	0 (0)
STI[c] test	120 (48.6)	103 (85.8)	7 (5.8)	11 (11.2)	1 (1.0)	3 (3.1)
Physician	178 (72.1)	161 (90.4)	5 (2.8)	7 (13.7)	0 (0)	3 (5.9)
Pharmacy	183 (74.1)	169 (92.3)	3 (1.6)	7 (5.9)	0 (0)	0 (0)
Free condoms	78 (31.6)	64 (82.1)	3 (3.8)	9 (7.5)	1 (0.8)	0 (0)

[a]Participants reported up to 2 socialization locations.

[b]Participants reported up to 3 sex locations.

[c]HIV: human immunodeficiency virus; STI: sexually transmitted infection.

[d]Among participants who do not report being HIV positive.

In unadjusted analyses, less than college education was associated with not reporting home location (P=.003) and sex locations (P=.05) (Figure 2). Nonwhite race was significantly associated with not reporting physician (P=.01) locations. Sex at the partner's house was significantly associated with not reporting the sex location (P=.001). No other bivariate associations were statistically significant.

In adjusted analyses, only 4 covariates were significantly associated with answering specific map questions (Figure 3). Nonwhite participants were less likely than white participants to locate a pharmacy (PR, 0.8; 95% CI, 0.7-1.0). Participants living in Georgia were more likely than participants living in Wisconsin to locate a primary care physician (PR, 1.3; 95% CI, 1.0-1.6). Participants reporting sex at their partner's home were less likely to indicate the sex location (PR, 0.8; 95% CI, 0.7-1.0). Similarly, participants with less than a college education were less likely to indicate a sex location than participants with a college degree (PR, 0.7; 95% CI, 0.5-1.0).

No other model-based associations between the covariates and answering specific map questions were statistically significant. For example, participants who reported CAI were no more likely to report sex locations (PR, 1.0; 95% CI, 0.9-1.1).

Data Accuracy

Of the 226 participants whose map-based home location could be assigned to a ZCTA, 187 (83%) placed the home location within the boundaries of the reported residential ZIP code. Of the 39 participants (17%) who placed a home location outside of the boundaries of the reported residential ZIP code, 29 placed the home location in an adjacent ZIP code, 2 reported post office box or institutional ZIP codes with a correct pin drop, and 8 placed the home location in a nonadjacent ZIP code. Reliability of residential location did not vary with urbanicity (P=.15).

Of the 1176 unique locations reported by the participants, the combination of the location type and the participant's text description permitted 575 locations (49%) to be identified. Of these, 278 text descriptions (48%) were a specific name (eg, Walgreens), 61 (11%) were a geographic area (eg, downtown, San Antonio), and 236 (41%) were a generic name (eg, doctor, pharmacy, hospital). Of the 61 locations identified as a geographic area, 53 (87%) were placed in the correct geographic area. Locations were not able to be identified because of a name that had meaning only to the participant (eg, home, work, guy 2's place, RLD).

Of all locations having a specific text description, the median distance between the participant-selected location and the location determined using the specific text description was 0.29 miles (interquartile range, 0.06-0.88). Of all locations having a generic text description, the median distance between the selected location and the location determined using the generic text description was 0.29 miles (IQR, 0.08-0.64). When stratified by location type, median distances between the selected location and location determined using the text descriptions were generally <one-third mile (Table 3). Although home and work have the highest median distances, very few locations could be identified based on the participant's text description.

Table 3. Distance in miles between selected location and location determined using any text description.

Location	Count	Median	IQR
Home	4	0.61	0.49-0.64
Work	9	0.77	0.57-2.88
Socialize	154	0.33	0.09-0.92
Doctor	141	0.19	0.05-0.65
Pharmacy	45	0.37	0.12-0.89
Sex	86	0.34	0.10-0.74
Condoms	31	0.08	0.01-0.50
HIV[a] test	19	0.22	0.01-0.52
STD test	24	0.22	0.13-0.49

[a]HIV: human immunodeficiency virus; STD: sexually transmitted disease.

Discussion

Principal Findings

In this paper, we examined the feasibility of collecting location-based data using a Web-based, map-based tool among an online convenience sample of MSM. Overall, participants were willing and able to use this tool to accurately indicate the requested locations, suggesting that this method of data collection is feasible, and results in complete, good quality data. In addition, for most locations, men who chose to not use the map tool were not significantly different from men who did use the tool with respect to demographic factors and HIV-related behaviors. The notable exception to this finding is that men were 20% less likely to report a sex location if that location was a partner's home, reflecting both confidentiality concerns and uncertainty in the exact location.

The lack of significant associations between the examined covariates and using the map tool has critical implications for the use and subsequent interpretation of these data. Analyses relying on these locations in similar Web-based populations will have minimal bias resulting from nonresponse to these questions, with respect to the covariates measured in this study, although bias may exist due to nonparticipation. A first key exception to this finding was the observed educational gradient in which participants with no college education were less likely to provide all requested locations and sex locations. Missing data among these individuals may especially be a concern in Web-based research, where MSM of color are more difficult to recruit [47].

A second key exception is the potential for bias in analyses using sex location when sex occurs at the partner's home (although a large majority still provided this location). Therefore, these missing data may bias analyses where either having sex at the partner's home or education is associated with both the exposure and outcome [48]. This finding may be critical for confounding by education because lower levels of education are frequently associated with locations and with poorer health outcomes.

Men who did not provide the requested locations were generally unwilling, rather than unable, to provide the locations. Even in an anonymous Web-based survey, privacy remained a concern among a small fraction of participants. Although most participants responded to these map questions, privacy concerns for these few individuals must be considered in the implementation and interpretation of future surveys. Providing participants with the opportunity to learn more about their data's security and reinforcing the acceptability of reporting approximate locations (eg, the nearest intersection) may help to assuage these concerns.

Similarly, participants' inability to provide these locations could also be addressed within the Web-based survey. This inability may stem from a lack of geographic knowledge or uncertainty in locations. Incorporating text search boxes to search for a given street name or emphasizing the acceptability of identifying an intersection or other landmark could potentially address this limitation. This recommendation could also reduce the observed educational gradient in responding to these questions.

As with all participant-reported data, reliability is an important concern. Despite asking numerous locations for which participants may not readily know an address, we found good agreement between the reported locations and other reported characteristics of those locations. These results are similar to the results of a prior validation of this tool for home and treatment locations among HIV-positive MSM [32]. Participants generally placed home locations within the correct ZIP code and placed other types of locations near the probable true location. This finding suggests that, although precise measures should be used with caution, within-person and between-person relative measures are likely appropriate.

Our findings with respect to answering specific questions contrast with past studies of broader Web-based survey participation. These studies found differential participation in Web-based surveys by demographic and health-related factors. Nonurban MSM have participated in Web-based surveys more than their urban counterparts [34,35]. In addition, individuals with a given medical condition are more likely to participate in research about that condition [36], suggesting that HIV-negative men could have been less likely to provide the requested data compared with men living with HIV.

Compared with previous validation studies [32], this analysis has expanded both the population and types of locations for which valid Web-based map data may be collected. We included MSM, independent of HIV status, from urban, suburban, and rural locations, not only large urban areas that are the typical geographic focus of much HIV research. We also included a wide variety of nonresidential locations that may be contextually important to the health of MSM.

As this study verified that these nonresidential location data can be collected from online samples of MSM, these locations may now be used to describe the activity spaces of MSM and to explore associations between nonresidential places and HIV-related behaviors among MSM. This Web-based tool will permit these location data to be collected using relatively low-resource methods that preserve participants' anonymity. The results of future analyses may allow us to better consider how differing contexts are associated with HIV risk and prevention. National surveillance data, which are based on residential locations, may be interpreted differently depending on the spatial variation in HIV-related behaviors. In addition, future analyses may permit interventions and policy to be geographically targeted using the locations of relevant behaviors, rather than residential locations.

Limitations

Despite the breadth of data being collected, this study does have limitations. First, the generalizability of this study may be limited. Our online convenience sample is likely not representative of MSM in Texas, Wisconsin, and Georgia. Our sample is less racially diverse, younger, and more educated compared with the general populations in these states. In addition, despite the breadth of HIV epidemiology, demographic, and contextual factors represented by these states, these MSM may not be representative of MSM across the United States. However, prior studies using venue-time–based sampling of MSM reported demographics similar to this study and to the Internet samples of MSM [37,47,49].

This analysis produced fully-adjusted measures of association for a large number of outcomes and their potential predictors. Consequently, some of these measures may be statistically significant due to type 2 error.

This analysis also used participant-reported ZIP codes as the basis for poverty and urbanicity measures. The use of areas to represent contextual variables may lead to misclassification, especially when using ZCTAs to represent ZIP codes [50,51]. The degree of this misclassification may be less in more urban areas [52-55], although this was not true in our predominantly urban sample. However, ZIP codes are a geographic measure that is readily accessible to participants and are therefore useful despite their limitations.

This study also was unable to validate all locations using a physical address. With our study's expansion to locations that include where individuals socialized and had sex, validation becomes more difficult as participants may not readily know addresses of these nonresidential locations. Consequently, data reliability could be assessed only using the methods we used. In addition, the text descriptions of these places were useful for only half of locations, limiting conclusions regarding reliability of the remaining half of locations. It is possible that the half of locations that could be validated may have favorably biased the calculated accuracy. Additional validation of geographic reliability may be the subject of future work.

Conclusions

Using a Web-based map tool, MSM participants were generally willing and able to indicate all requested locations. Critically, although most MSM reported sex locations, these locations were reported less frequently than all other locations. Consequently, within this Web-based setting and MSM population (and with careful consideration of the potential biases associated with Web-based research in this population), this method of data collection is feasible, resulting in highly complete, good quality location data.

Conflicts of Interest

None declared.

References

1. Cummins S, Curtis S, Diez-Roux AV, Macintyre S. Understanding and representing 'place' in health research: a relational approach. Soc Sci Med 2007 Nov;65(9):1825-1838. [doi: 10.1016/j.socscimed.2007.05.036] [Medline: 17706331]
2. Centers for Disease Control and Prevention. Atlanta, GA; 2015 Feb. HIV surveillance report: Diagnoses of HIV infection in the United States and dependent areas, 2013 URL: http://www.cdc.gov/hiv/library/reports/surveillance/ [accessed 2016-05-24] [WebCite Cache ID 6hkL6VTQv]
3. Buot MG, Docena JP, Ratemo BK, Bittner MJ, Burlew JT, Nuritdinov AR, et al. Beyond race and place: distal sociological determinants of HIV disparities. PLoS One 2014;9(4):e91711 [FREE Full text] [doi: 10.1371/journal.pone.0091711] [Medline: 24743728]

XSL•FO

RenderX

4. Frye V, Latka MH, Koblin B, Halkitis PN, Putnam S, Galea S, et al. The urban environment and sexual risk behavior among men who have sex with men. J Urban Health 2006 Mar;83(2):308-324 [FREE Full text] [doi: 10.1007/s11524-006-9033-x] [Medline: 16736379]

5. Hixson BA, Omer SB, del RC, Frew PM. Spatial clustering of HIV prevalence in Atlanta, Georgia and population characteristics associated with case concentrations. J Urban Health 2011 Feb;88(1):129-141 [FREE Full text] [doi: 10.1007/s11524-010-9510-0] [Medline: 21249526]

6. Vaughan AS, Rosenberg ES, Sullivan PS. Spatial relationships between gay stigma, poverty, and HIV infection among black and white men who have sex with men in Atlanta. AIDS Res Hum Retroviruses 2014;30(8):740-741. [Medline: 25013969]

7. Laraque F, Mavronicolas HA, Robertson MM, Gortakowski HW, Terzian AS. Disparities in community viral load among HIV-infected persons in New York City. AIDS 2013 Aug 24;27(13):2129-2139. [doi: 10.1097/QAD.0b013e328360f619] [Medline: 24384591]

8. Krieger N, Chen JT, Waterman PD, Rehkopf DH, Subramanian SV. Painting a truer picture of US socioeconomic and racial/ethnic health inequalities: the Public Health Disparities Geocoding Project. Am J Public Health 2005 Feb;95(2):312-323. [doi: 10.2105/AJPH.2003.032482] [Medline: 15671470]

9. Bowleg L, Neilands TB, Tabb LP, Burkholder GJ, Malebranche DJ, Tschann JM. Neighborhood context and Black heterosexual men's sexual HIV risk behaviors. AIDS Behav 2014 Nov;18(11):2207-2218. [doi: 10.1007/s10461-014-0803-2] [Medline: 24906531]

10. Jeffries WL, Marks G, Lauby J, Murrill CS, Millett GA. Homophobia is associated with sexual behavior that increases risk of acquiring and transmitting HIV infection among black men who have sex with men. AIDS Behav 2013 May;17(4):1442-1453. [doi: 10.1007/s10461-012-0189-y] [Medline: 22569942]

11. Ayala G, Bingham T, Kim J, Wheeler DP, Millett GA. Modeling the impact of social discrimination and financial hardship on the sexual risk of HIV among Latino and Black men who have sex with men. Am J Public Health 2012 May;102(S2):S242-S249. [doi: 10.2105/AJPH.2011.300641] [Medline: 22401516]

12. Carpiano RM, Kelly BC, Easterbrook A, Parsons JT. Community and drug use among gay men: the role of neighborhoods and networks. J Health Soc Behav 2011 Mar;52(1):74-90. [doi: 10.1177/0022146510395026] [Medline: 21362613]

13. Buttram ME, Kurtz SP. Risk and protective factors associated with gay neighborhood residence. Am J Mens Health 2013;7(2):110-118 [FREE Full text] [doi: 10.1177/1557988312458793] [Medline: 22948299]

14. Koblin BA, Egan JE, Rundle A, Quinn J, Tieu H, Cerdá M, et al. Methods to measure the impact of home, social, and sexual neighborhoods of urban gay, bisexual, and other men who have sex with men. PLoS One 2013;8(10):e75878 [FREE Full text] [doi: 10.1371/journal.pone.0075878] [Medline: 24146785]

15. Duncan DT, Kapadia F, Halkitis PN. Examination of spatial polygamy among young gay, bisexual, and other men who have sex with men in New York City: the P18 cohort study. Int J Environ Res Public Health 2014;11(9):8962-8983 [FREE Full text] [doi: 10.3390/ijerph110908962] [Medline: 25170685]

16. Verghese A, Berk SL, Sarubbi F. Urbs in rure: human immunodeficiency virus infection in rural Tennessee. J Infect Dis 1989 Dec;160(6):1051-1055. [Medline: 2584752]

17. Lansky A, Nakashima AK, Diaz T, Fann SA, Conti L, Herr M, et al. Human immunodeficiency virus infection in rural areas and small cities of the southeast: contributions of migration and behavior. J Rural Health 2000;16(1):20-30. [Medline: 10916312]

18. Tobin KE, Latkin CA, Curriero FC. An examination of places where African American men who have sex with men (MSM) use drugs/drink alcohol: a focus on social and spatial characteristics. Int J Drug Policy 2014 May;25(3):591-597 [FREE Full text] [doi: 10.1016/j.drugpo.2013.12.006] [Medline: 24484732]

19. Oster AM, Wejnert C, Mena LA, Elmore K, Fisher H, Heffelfinger JD. Network analysis among HIV-infected young black men who have sex with men demonstrates high connectedness around few venues. Sex Transm Dis 2013 Mar;40(3):206-212. [doi: 10.1097/OLQ.0b013e3182840373] [Medline: 23403601]

20. Krieger N. Theories for social epidemiology in the 21st century: an ecosocial perspective. Int J Epidemiol 2001 Aug;30(4):668-677 [FREE Full text] [Medline: 11511581]

21. Adimora AA, Schoenbach VJ. Social context, sexual networks, and racial disparities in rates of sexually transmitted infections. J Infect Dis 2005 Feb 1;191(S1):S115-S122 [FREE Full text] [doi: 10.1086/425280] [Medline: 15627221]

22. Latkin CA, German D, Vlahov D, Galea S. Neighborhoods and HIV: a social ecological approach to prevention and care. Am Psychol 2013;68(4):210-224 [FREE Full text] [doi: 10.1037/a0032704] [Medline: 23688089]

23. Baral S, Logie CH, Grosso A, Wirtz AL, Beyrer C. Modified social ecological model: a tool to guide the assessment of the risks and risk contexts of HIV epidemics. BMC Public Health 2013;13(1):482 [FREE Full text] [doi: 10.1186/1471-2458-13-482] [Medline: 23679953]

24. Rothenberg R. Maintenance of endemicity in urban environments: a hypothesis linking risk, network structure and geography. Sex Transm Infect 2007 Feb;83(1):10-15 [FREE Full text] [doi: 10.1136/sti.2006.017269] [Medline: 17283360]

25. Matthews SA, Yang T. Spatial Polygamy and Contextual Exposures (SPACEs): Promoting Activity Space Approaches in Research on Place and Health. Am Behav Sci 2013 Aug 1;57(8):1057-1081 [FREE Full text] [doi: 10.1177/0002764213487345] [Medline: 24707055]

26. Zenk SN, Schulz AJ, Matthews SA, Odoms-Young A, Wilbur J, Wegrzyn L, et al. Activity space environment and dietary and physical activity behaviors: a pilot study. Health Place 2011 Sep;17(5):1150-1161 [FREE Full text] [doi: 10.1016/j.healthplace.2011.05.001] [Medline: 21696995]

27. Boruff BJ, Nathan A, Nijënstein S. Using GPS technology to (re)-examine operational definitions of 'neighbourhood' in place-based health research. Int J Health Geogr 2012;11:22 [FREE Full text] [doi: 10.1186/1476-072X-11-22] [Medline: 22738807]

28. Chaix B, Méline J, Duncan S, Merrien C, Karusisi N, Perchoux C, et al. GPS tracking in neighborhood and health studies: a step forward for environmental exposure assessment, a step backward for causal inference? Health Place 2013 May;21:46-51. [doi: 10.1016/j.healthplace.2013.01.003] [Medline: 23425661]

29. Rainham D, McDowell I, Krewski D, Sawada M. Conceptualizing the healthscape: contributions of time geography, location technologies and spatial ecology to place and health research. Soc Sci Med 2010 Mar;70(5):668-676. [doi: 10.1016/j.socscimed.2009.10.035] [Medline: 19963310]

30. Delaney KP, Kramer MR, Waller LA, Flanders WD, Sullivan PS. Using a geolocation social networking application to calculate the population density of sex-seeking gay men for research and prevention services. J Med Internet Res 2014;16(11):e249 [FREE Full text] [doi: 10.2196/jmir.3523] [Medline: 25406722]

31. Duncan DT, Kapadia F, Regan SD, Goedel WC, Levy MD, Barton SC, et al. Feasibility and Acceptability of Global Positioning System (GPS) Methods to Study the Spatial Contexts of Substance Use and Sexual Risk Behaviors among Young Men Who Have Sex with Men in New York City: A P18 Cohort Sub-Study. PLoS One 2016 Feb;11(2):e0147520 [FREE Full text] [doi: 10.1371/journal.pone.0147520] [Medline: 26918766]

32. Dasgupta S, Vaughan AS, Kramer MR, Sanchez TH, Sullivan PS. Use of a Google Map Tool Embedded in an Internet Survey Instrument: Is it a Valid and Reliable Alternative to Geocoded Address Data? JMIR Res Protoc 2014;3(2):e24 [FREE Full text] [doi: 10.2196/resprot.2946] [Medline: 24726954]

33. van Gelder MHJ, 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]

34. Sanchez TH, Sineath RC, Kahle EM, Tregear SJ, Sullivan PS. The annual American Men's Internet Survey of behaviors of men who have sex with men in the United States: Protocol and key indicators report 2013. JMIR Public Heal Surveill 2015;1(1):e3. [doi: 10.2196/publichealth.4314]

35. Horvath KJ, Rosser B R Simon, Remafedi G. Sexual risk taking among young internet-using men who have sex with men. Am J Public Health 2008 Jun;98(6):1059-1067. [doi: 10.2105/AJPH.2007.111070] [Medline: 18445804]

36. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007 Sep;17(9):643-653. [doi: 10.1016/j.annepidem.2007.03.013] [Medline: 17553702]

37. Hernandez-Romieu AC, Sullivan PS, Sanchez TH, Kelley CF, Peterson JL, Del RC, et al. The comparability of men who have sex with men recruited from venue-time-space sampling and facebook: a cohort study. JMIR Res Protoc 2014;3(3):e37 [FREE Full text] [doi: 10.2196/resprot.3342] [Medline: 25048694]

38. Centers for Disease Control and Prevention. 2015. HIV Among Youth Internet URL: http://www.cdc.gov/hiv/risk/age/youth/index.html [accessed 2016-02-26] [WebCite Cache ID 6faaq4WDM]

39. Centers for Disease Control and Prevention. 2015. HIV Among People Aged 50 and Older Internet URL: http://www.cdc.gov/hiv/risk/age/olderamericans/index.html [accessed 2016-02-26] [WebCite Cache ID 6fafddygy]

40. U.S. Census Bureau. 1995. Poverty Areas Internet URL: http://www.census.gov/population/socdemo/statbriefs/povarea.html [accessed 2016-02-26] [WebCite Cache ID 6fafhu529]

41. Ingram DD, Franco SJ. 2013 NCHS Urban-Rural Classification Scheme for Counties. Vital Health Stat 2 2014 Apr(166):1-73 [FREE Full text] [Medline: 24776070]

42. Graubard BI, Korn EL. Predictive margins with survey data. Biometrics 1999 Jun;55(2):652-659. [Medline: 11318229]

43. Bieler GS, Brown GG, Williams RL, Brogan DJ. Estimating model-adjusted risks, risk differences, and risk ratios from complex survey data. Am J Epidemiol 2010 Mar 1;171(5):618-623 [FREE Full text] [doi: 10.1093/aje/kwp440] [Medline: 20133516]

44. McNutt L, Wu C, Xue X, Hafner JP. Estimating the relative risk in cohort studies and clinical trials of common outcomes. Am J Epidemiol 2003 May 15;157(10):940-943 [FREE Full text] [Medline: 12746247]

45. Williamson T, Eliasziw M, Fick GH. Log-binomial models: exploring failed convergence. Emerg Themes Epidemiol 2013;10(1):14 [FREE Full text] [doi: 10.1186/1742-7622-10-14] [Medline: 24330636]

46. R-project. Vienna, Austria: R Foundation for Statistical Computing; 2015. A Language and Environment for Statistical Computing URL: https://www.r-project.org/ [accessed 2016-05-24] [WebCite Cache ID 6hkMotC0u]

47. Sullivan PS, Khosropour CM, Luisi N, Amsden M, Coggia T, Wingood GM, et al. Bias in online recruitment and retention of racial and ethnic minority men who have sex with men. J Med Internet Res 2011;13(2):e38 [FREE Full text] [doi: 10.2196/jmir.1797] [Medline: 21571632]

48. Westreich D. Berkson's bias, selection bias, and missing data. Epidemiology 2012 Jan;23(1):159-164 [FREE Full text] [doi: 10.1097/EDE.0b013e31823b6296] [Medline: 22081062]

XSL•FO

RenderX

49. Sullivan PS, Peterson J, Rosenberg ES, Kelley CF, Cooper H, Vaughan A, et al. Understanding racial HIV/STI disparities in black and white men who have sex with men: a multilevel approach. PLoS One 2014;9(3):e90514 [FREE Full text] [doi: 10.1371/journal.pone.0090514] [Medline: 24608176]

50. Krieger N, Waterman P, Chen JT, Soobader M, Subramanian SV, Carson R. Zip code caveat: bias due to spatiotemporal mismatches between zip codes and US census-defined geographic areas--the Public Health Disparities Geocoding Project. Am J Public Health 2002 Jul;92(7):1100-1102. [Medline: 12084688]

51. Duncan DT, Kawachi I, Subramanian SV, Aldstadt J, Melly SJ, Williams DR. Examination of how neighborhood definition influences measurements of youths' access to tobacco retailers: a methodological note on spatial misclassification. Am J Epidemiol 2014 Feb 1;179(3):373-381 [FREE Full text] [doi: 10.1093/aje/kwt251] [Medline: 24148710]

52. Healy MA, Gilliland JA. Quantifying the magnitude of environmental exposure misclassification when using imprecise address proxies in public health research. Spat Spatiotemporal Epidemiol 2012 Apr;3(1):55-67. [doi: 10.1016/j.sste.2012.02.006] [Medline: 22469491]

53. Boone JE, Gordon-Larsen P, Stewart JD, Popkin BM. Validation of a GIS facilities database: quantification and implications of error. Ann Epidemiol 2008 May;18(5):371-377 [FREE Full text] [doi: 10.1016/j.annepidem.2007.11.008] [Medline: 18261922]

54. Bonner MR, Han D, Nie J, Rogerson P, Vena JE, Freudenheim JL. Positional accuracy of geocoded addresses in epidemiologic research. Epidemiology 2003 Jul;14(4):408-412. [doi: 10.1097/01.EDE.0000073121.63254.c5] [Medline: 12843763]

55. Vieira VM, Howard GJ, Gallagher LG, Fletcher T. Geocoding rural addresses in a community contaminated by PFOA: a comparison of methods. Environ Health 2010;9:18 [FREE Full text] [doi: 10.1186/1476-069X-9-18] [Medline: 20406495]

Abbreviations

CAI: condomless anal intercourse
GPS: global positioning systems
HIV: human immunodeficiency virus
MSM: men who have sex with men
POR: prevalence odds ratio
PR: prevalence ratio
STI: sexually transmitted infection
ZCTA: zip code tabulation area

XSL•FO
RenderX

Original Paper

Using Facebook to Reach People Who Experience Auditory Hallucinations

Benjamin Sage Crosier[1], PhD; Rachel Marie Brian[2], MPH; Dror Ben-Zeev[2], PhD

[1]Center for Technology and Behavioral Health, Department of Biomedical Data Science, Dartmouth College, Lebanon, NH, United States

[2]mHealth for Mental Health Program, Dartmouth Psychiatric Research Center, Department of Psychiatry, Dartmouth College, Lebanon, NH, United States

Corresponding Author:
Benjamin Sage Crosier, PhD
Center for Technology and Behavioral Health
Department of Biomedical Data Science
Dartmouth College
Suite 300
46 Centerra Parkway
Lebanon, NH, 03766
United States
Phone: 1 5185965379
Fax: 1 603 448 3976
Email: benjamin.crosier@dartmouth.edu

Abstract

Background: Auditory hallucinations (eg, hearing voices) are relatively common and underreported false sensory experiences that may produce distress and impairment. A large proportion of those who experience auditory hallucinations go unidentified and untreated. Traditional engagement methods oftentimes fall short in reaching the diverse population of people who experience auditory hallucinations.

Objective: The objective of this proof-of-concept study was to examine the viability of leveraging Web-based social media as a method of engaging people who experience auditory hallucinations and to evaluate their attitudes toward using social media platforms as a resource for Web-based support and technology-based treatment.

Methods: We used Facebook advertisements to recruit individuals who experience auditory hallucinations to complete an 18-item Web-based survey focused on issues related to auditory hallucinations and technology use in American adults. We systematically tested multiple elements of the advertisement and survey layout including image selection, survey pagination, question ordering, and advertising targeting strategy. Each element was evaluated sequentially and the most cost-effective strategy was implemented in the subsequent steps, eventually deriving an optimized approach. Three open-ended question responses were analyzed using conventional inductive content analysis. Coded responses were quantified into binary codes, and frequencies were then calculated.

Results: Recruitment netted N=264 total sample over a 6-week period. Ninety-seven participants fully completed all measures at a total cost of $8.14 per participant across testing phases. Systematic adjustments to advertisement design, survey layout, and targeting strategies improved data quality and cost efficiency. People were willing to provide information on what triggered their auditory hallucinations along with strategies they use to cope, as well as provide suggestions to others who experience auditory hallucinations. Women, people who use mobile phones, and those experiencing more distress, were reportedly more open to using Facebook as a support and/or therapeutic tool in the future.

Conclusions: Facebook advertisements can be used to recruit research participants who experience auditory hallucinations quickly and in a cost-effective manner. Most (58%) Web-based respondents are open to Facebook-based support and treatment and are willing to describe their subjective experiences with auditory hallucinations.

XSL•FO

RenderX

Introduction

Seven percent of the population has experienced auditory hallucinations (eg, hearing voices) [1]. Many individuals experience auditory hallucinations in the context of a serious mental illness (SMI). Approximately 40% of people with SMI go undetected and untreated by traditional mental health care systems [2], a shortcoming exacerbated by environmental hurdles and poverty [3]. Many people with SMI report wanting to deal with their problems on their own rather than seek traditional treatment [2]. Novel outreach methods have the potential to reach individuals who experience auditory hallucinations and who are unwilling or unable to use traditional mental health services.

Traditional engagement methods for research through face-to-face contact at brick and mortar institutions, such as clinics or academic institutions, face numerous geographical and practical barriers (eg, stigmatization associated with going into a clinic, transportation). In contrast, social media offers an unparalleled opportunity to engage hard-to-reach populations, as most American adults (74%) use social networking sites [4], including those with a variety of health problems [5]. Facebook is the most successful Web-based social network to date, with nearly a billion and a half frequent users [6]. Other popular services like Instagram (400 million active users) and Snapchat (approaching 200 million active users), along with a plethora of smaller specialized networks, may offer additional opportunities for outreach [7]. These diverse user bases can be reached through advertising systems typically used for marketing through the promotion of posts, pages, apps, and Websites. Treatment information, Web-based surveys, and other recruitment materials can be distributed to vast audiences on modest budgets and within flexible timeframes in this advertisement ecosystem. Furthermore, advertising platforms have filters that allow for accurate targeting of personal characteristics, along with systems for real-time tracking of advertisement performance [8].

Facebook advertisements could be used as an effective recruitment tool for mental health studies, excelling in accessing hard-to-reach populations. Previous research has shown that most young people (94%) with SMI use social media [9], and most people with SMI use mobile technologies for communication and the Internet access [10]. Facebook has also been used to successfully recruit Veteran's for mental health research [11]. Kosinski and colleagues offer a thorough general review of using Facebook as a research tool in any social science, highlighting this approach's ability to rapidly collect data on millions of participants efficiently [12]. However, little is known about best approaches to reach those with SMI through the use of social media.

Previous research has demonstrated that overall survey length, topic, question length, sponsoring institution (eg, private industry, government, academia), all affect response rates of traditional surveys, but little work has been done to test these findings to Web-based surveys [13]. The success of using social media as an outreach or recruitment tool is partly determined by maximizing each step of the data collection process.

Advertisements need to be tested for relative effectiveness, honing in on materials that entice the most social media users to take action for the lowest cost to researchers. This approach is common practice in marketing. Advertisers commonly use "split tests" or "A/B testing" to hone their marketing strategies, on and off the Web [14]. Although general design principles guide advertisement design, the effectiveness of images is highly dependent on the nature of the audience and topic. Standard advertising practices suggest testing multiple advertisement layouts to determine efficiency, and the design of content is largely dependent on expert knowledge, in this case, auditory hallucinations.

Surveys need to be designed in a way that promotes conscientious and full participation at each step, discouraging incomplete responses and dropout. This research systematically tested these elements, arriving at an optimized advertisement and survey for people who experience auditory hallucinations. We detail the steps in this process and provide a cost analysis. Furthermore, we include insights gained from the survey itself regarding social media attitudes related to Web-based support and treatment. Specifically, we assess whether respondents would be willing to use social media to connect to peers experiencing similar problems or to clinicians to engage in services. We also present thematic frequencies from responses given by participants regarding what makes their hallucinations worse, ways in which they cope, and advice they would give to others who also experience auditory hallucinations.

Methods

The Committee for Protection of Human Subjects at Dartmouth College approved this study. Our team ran a series of Facebook advertisements in 6 phases (see Figure 1) to optimize response rate and data quality by distributing an 18-item survey (see Multimedia Appendix 1) between June 9 and July 17, 2015. We first piloted a set of candidate advertisements to ensure acceptability within Facebook's advertising system. We then tested each of the images in 2 layout styles (single image vs a multiple image "carousel") for relative performance. Next, we evaluated survey layout, maximizing data quality. We tested survey pagination, question type (multiple choice vs free response) ordering, and advertisement-targeting strategies. We concluded with a data collection run using optimized advertisement and survey designs to obtain a sufficient sample size.

Survey items captured demographics, symptom history, technology use, and attitudes, and solicited participant's own experiences with auditory hallucinations and advice to help others (see Multimedia Appendix 1). To completely anonymize the survey, all automatic data collection features were turned off (eg, IP logging), and no identifying information was collected. Participants were eligible to participate if they were aged older than 18 years and confirmed that they experience auditory hallucinations, which was confirmed via survey questions (see Figure 2 for a screenshot of the survey platform hosted in Survey Monkey). Data from incomplete responses were logged by Survey Monkey, making it possible to determine

XSL•FO

RenderX

click through rates. All advertisement distribution data, including cost, were monitored with Facebook's advertising system.

Three open-ended questions were included in the Web-based survey, each having 3 fields for participants to provide responses. Questions included were *(1) What are 3 things that make your voices worse? (2) What are your top 3 methods for coping with voices? (3) What are 3 things people should keep in mind if they don't want voices to control their lives?* Participant responses to the 3 open-ended questions were analyzed using conventional inductive content analysis [15]. Two research assistants (RW and GJ) with direct oversight from author (RB), independently read responses from each open-ended question to identify recurring regularities among entries to establish a preliminary list of codes [16]. Through group discussion and review, the induced codes were organized into categories to create the codebook [17,18]. Each research assistant then coded responses separately using this codebook.

Once coding was complete, the 2 coders met to discuss their coded responses. Inter-rater reliability was calculated with 97.6% agreement. The rationale for why differently coded items were chosen was discussed until consensus was reached. Author RB then reviewed all coded content. When coding was finished, coded responses were *quantified* by converting responses to a binary format that was then used to calculate thematic frequencies [16]. Because there were 3 entries possible for each question, and to obtain a clear understanding of the unique responses given by each participant, similarly coded multiple responses given by a single participant were entered once into the database. For example, if an individual were to give the same response for each of the 3 entry fields for a question (eg, "ignore the voices," "do not listen to the voices," and "ignore") these would be entered into the binary database for that participant and variable once. Thematic frequencies were then calculated from this database.

Figure 1. Engagement Optimization Phases.

Figure 2. Screenshot of the survey platform.

Results

All data collection phases ($2150 total advertisement budget) collective garnered N=264 responses resulting in a final sample size N=97 participants that provided complete data. The cost per complete participant (N=97) in the final optimized stage was $8.14. The sample was 62% female with an average age of 28.91, standard deviation (SD)=14.50. The sample was composed of 57% white, 21% multiple race, 15% black, 4% American Indian or Alaskan Indian, and 3% Asian participants. A majority (82%) of the sample had at least a high school

diploma or had obtained a General Educational Development (GED) certification. Participants were previously hospitalized for mental health problems M=1.97 (SD=4.57) times during their lifetime. Each phase of outreach optimization is sequentially described in the following sections. All quantitative data were analyzed with the R programming language. We include a series of tables to provide descriptives results on items not reported in main analyses. Tables 1 and 2 include self-reported information on diagnoses and symptom descriptions respectively. Table 3 covers treatment history, and Table 4 details technology use.

Table 1. Self-reported information on diagnoses.

Psychiatric diagnoses	Frequency	Percent
Schizophrenia	15	9.1
Schizoaffective	9	5.5
Bipolar	39	23.6
Depression	68	41.2
PTSD	29	17.6

Table 2. Frequency tables for mental health–related survey items.

Item	Response	Frequency	Percent
"How often do you hear voices?"			
	Not currently/less than once	39	24.4
	At least once a week	37	23.1
	At least once a day	39	24.4
	At least once an hour	10	6.3
	Almost continuously/continuously	35	21.9
"How long have you been hearing voices?"			
	<1 Month	4	2.5
	1-6 Months	12	7.5
	7-11 Months	15	9.4
	1-5 Years	38	23.9
	>5 Years	90	56.6
"When they occur, how intense is your distress from the voices?"			
	Not distressing	40	25.3
	Slightly distressing	46	29.1
	Moderately distressing	32	20.3
	Very distressing	28	17.7

Table 3. Treatment history.[a]

Response	Frequency	Percent
Medications	44	26.7
Group therapy	21	12.7
Individual therapy	39	23.6
Peer support	12	12.7
Mindfulness	32	19.4

[a]Participants also had the option of filling in a free-text "other" option. Two respondents listed "nothing" and the following items were each listed once: "control my surroundings," "intensive inpatient and outpatient program," "listening to them to see if I can help," "marijuana," "meditations," "music: the voices sing with me and are not tone deaf," "myself," "none," "not really, my other personalities dobt (*sic*) really let me," "psilocybin mushrooms."

Table 4. Technology use–related items.

	Response	Frequency	Percent
Basic cell phone use			
	Do not have	53	42.7
	Every day	54	43.5
	4-6 days per week	5	3.2
	1 or less days per week	3	2.4
	Never use	10	8.1
Mobile phone use			
	Do not have	10	8.1
	Every day	104	83.9
	4-6 days per week	4	3.2
	1 or less days per week	2	1.6
	Never use	4	3.2
Tablet use			
	Do not have	63	52.5
	Every day	16	13.3
	4-6 days per week	3	2.5
	2-3 days per week	9	7.5
	1 or less days per week	13	10.8
	Never use	16	13.3
Wearable tech use			
	Do not have	78	47.3
	Every day	6	3.6
	4-6 days per week	2	1.2
	2-3 days per week	2	1.2
	1 or less days per week	2	1.2
	Never use	29	24.4
Email use			
	Do not have	8	4.8
	Every day	45	27.3
	4-6 days per week	15	9.1
	2-3 days per week	22	13.3
	1 or less days per week	23	13.9
	Never use	8	4.8
Social media use			
	Do not have	5	3.9
	Every day	101	79.5
	4-6 days per week	13	10.2
	2-3 days per week	5	3.9
	1 or less days per week	2	1.6
	Never use	1	0.8

Phase 1

The initial phase of advertisement lasted 24-hours and ran on a budget of $50. This phase verified the acceptability of advertisement materials, should a particular image be deemed unacceptable by Facebook. Seven advertisement images and 2 advertisement layouts (single image vs a multiple image "carousel") were tested, with 2 images not meeting Facebook's standards because more than 20% of the image area was occupied by text. These images were altered to meet the 20% requirement. The baseline cost or the ratio of the number of times the advertisements were presented to the number of times it was clicked on, was then evaluated. See Table 5 for a breakdown of cost-per-click descriptives by phase. Table 6 provides data on cost per participant from the phases it was available, as well as a discussion of this availability in the Conclusion section.

Phase 2

The second phase ran over a 1-week period and had a budget of $350. This phase evaluated the relative cost-effectiveness of advertisement materials by simultaneously running each of the 7 single image advertisements and a version of the carousel-type advertisement (each of the 5 images in this advertisement were populated by the most cost-effective images in the image pilot; see Figure 3).

Phase 3

This 1-week phase had a budget of $350. The optimized advertisement was used to test the influence of survey layout on response rate and data quality. This phase revealed no differences between grouping multiple questions per page versus having only one question on each page.

Phase 4

A subsequent 1-week phase with a $350 budget revealed that data quality was increased by placing open-ended free response questions at the end of the survey rather than at the beginning.

Phase 5

This 1-week phase used a budget of $350 to see if advertisement targeting could improve response rates. Facebook allows for the targeting of specific "interests." These are calculated by Facebook with factors varying from demographics to purchase history, profile information, wall posts, news feed, and other communications (propriety method from Facebook). We selected a collection of promental health interests ranging from general terms (eg, schizophrenia) to Web-based support groups (see Textbox 1 for a list of each interest). This targeted approach outperformed an open approach with no filters.

Table 5. Cost per click and engagement by advertising phase.

Phase	Version	Reach	Clicks	CPC[a]
1: Image piloting	N/A	13,292	138	$0.36
2: Image efficiency	N/A	54,134	1,717	$0.20
3: Pagination	Multiple pages	1047	34,411	$0.17
	One page	1085	30,035	$0.17
4: Question type ordering	Multiple choice first	1277	35,644	$0.17
	Free response first	1175	32,494	$0.15
5: Targeting	Open	1379	31,095	$0.15
	Targeted	919	31,092	$0.19
6: Open data collection	N/A	112,122	3302	$0.20

[a]CPC: cost per click.

Table 6. Cost per participant (partial and complete responses together) by advertising phase.[a]

Phase	Version	N	Cost
1: Pagination	Multiple pages		$11.67
	One page		$9.72
2: Question type ordering	Multiple choice first		$9.72
	Free response first		$7.61
3: Open data collection	N/A		$6.03

[a]Cost data only available for phases that manipulated survey-related changes and not advertisement-related changes. This information was also available for the final open data collection phase (VI). Please see the conclusion section for a discussion of this limitation. (CPP=cost per participant).

Figure 3. The final optimized advertisement. This advertisement was composed of 5 scrolling images that automatically rotated through a "carousel." Five different perspectives of the advertisement are depicted to display each image. Facebook automatically optimizes image order for cost efficiency. Users can manually scroll through images by clicking right or left arrows.
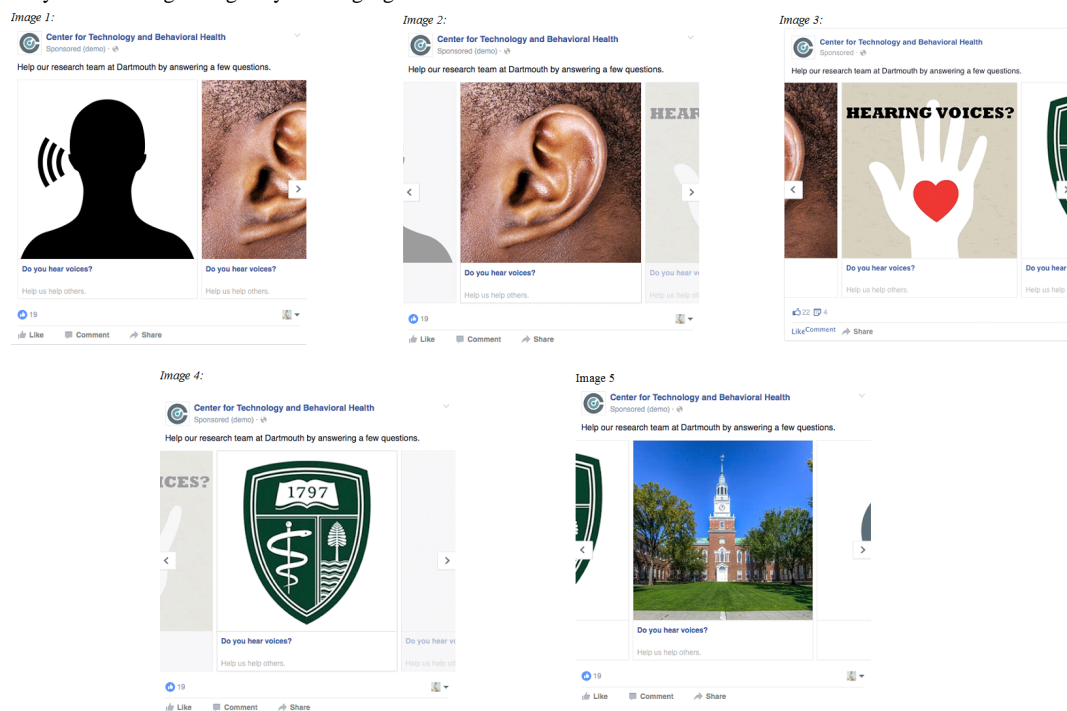
Textbox 1. Facebook "Interests" used for advertisement targeting.

Interests
Auditory hallucination
Bipolar disorder awareness
Delusional disorder
Depression and bipolar support alliance
Hallucination
Mental health
Paranoid schizophrenia
Psychiatry
Psychosis
Psychotherapy
Schizoaffective, depression, and bipolar support alliance
Schizoid personality disorder
Schizophrenia and smoking
Schizophrenia awareness
Schizophrenia research
Schizotypal personality disorder
Treatment of bipolar disorder

Phase 6

A final 2-week data collection phase had a total budget of $700. This phase did not perform any optimization test but was instead used to collect a sufficient sample size of N=97 participants who provided fully completed responses.

Next, we used 2 multiple regression models that were specified a priori to examine the predictors of interest in using Facebook as a mobile mental health tool. We examined 2 outcome

variables: interest in using Facebook to connect to similar others and interest in using Facebook to connect to clinicians. We included an identical set of predictors in each model that captured the level of distress from experiencing auditory hallucinations and technology use in each model, and controlled for demographics. Mobile phone use was predictive of an interest in using Facebook to connect with similar others (b=0.35, t(114)=2.18, P=.03). Females (b=−0.60, t(108)=−2.12, P=.04) and those who experienced more distress from hearing auditory hallucinations (b=0.30, t(135)=3.00, P=.003) expressed more interest in using Facebook to connect to clinicians.

Qualitative Findings

Ninety-seven participants gave at least one response for one or more of the open-ended questions for qualitative analysis. Two participants were younger than 18 years and were removed from the dataset, along with a participant whose entries seemed disconnected from the questions asked (e.g., responding *"I am a wiccan"* as advice for others to help them with their auditory hallucinations, responding *"Marshall Law"* as something that makes their hallucinations worse). This gave a total of 94 eligible participants who gave at least one comprehendible open-ended response.

Auditory Hallucination Triggers

Out of the 90 people who responded, 209 unique entries were analyzed. Close to a third (31%) of responses related to some form of emotional distress (stress, anxiety, depression, or anger) as sources of symptom exacerbation. Social environments, such as being alone, situations involving other people, or conflict with others represented 23% of responses. Environmental factors (including noise, silence, light/darkness, and time of day) made up 15% of responses. Physical states, such as lack of sleep, excessive sleep, fever, and menstrual cycle accounted for 13% of unique responses. Seven percent of responses included alcohol, marijuana, or illicit substances. Mental states, such as overthinking/overfocusing or "zoning out" accounted for 5% of responses. Three percent thought that paying attention to auditory hallucinations and internalized stigmatization associated with auditory hallucinations (eg, *"Telling others about them," "Thinking people have a reason to dislike me."*) made their hallucinations worse. Three percent reported not knowing what made their hallucinations worse.

Ways Participants Cope With Auditory Hallucinations

One-hundred ninety eight codes were analyzed from 89 participants who gave responses. Most coping methods included strategies that could be used in the moment, including distraction techniques, ways of rationalizing hallucinations, and self-medication. Distraction techniques accounted for 58% of responses and included behavioral strategies such as listening to music, singing, socializing, relaxation techniques, sleep, exercise, and reading. Choosing to directly engage or disengage with the hallucinations, accepting them, rationalizing who/what they are, talking with or listening to hallucinations, or ignoring hallucinations accounted for 29% of responses. Five percent cope with some form of self-medication (alcohol, marijuana, or illicit substances), 3% cope with prescription medication,

2% use religion (prayer), and 3% did not know or reported that nothing makes their hallucinations better.

Advice Given to Help Others With Auditory Hallucinations

One-hundred seventy unique codes were analyzed from 80 participants who gave responses. Most responses (30%) dealt with understanding auditory hallucinations. This included the way one rationalizes who/what they are, the control one has over their auditory hallucinations, engaging with auditory hallucinations, acceptance of them or learning more about why they happen. Twenty-two percent of responses included finding distraction from the auditory hallucinations (ie, listening to music, relaxation techniques) or ignoring auditory hallucinations altogether. Thirteen percent included statements of encouragement such as, *"Remember that it could be worse"* and *"Stay strong."* Nine percent of responses included suggestions to seek out mental health services or prescription medication. Another 9% listed that there was no help or that they did not know of any advice to give. Six percent of responses mentioned socializing, and another 6% contained religious statements (mostly about prayer). A small group advised the use of marijuana (2%) and 2% also listed other behavioral strategies (eat something, watch television, and so forth).

Conclusion

We present results from a Web-based survey that used Facebook advertisements as a data collection strategy, suggesting that social media is a viable approach to reach people who experience auditory hallucinations. This scalable and efficient method can be used to tap a hard to reach population—a strength that is particularly relevant in reaching people who face economic or practical barriers accessing in-person services (eg, geographical) or are turned off by other outreach strategies that are plagued by the hazards of stigmatization.

Outside of recruitment and collection optimization, the exploratory survey provided insight surrounding personal experiences of those who experience auditory hallucinations and technology use. Regarding recruitment, we were able to obtain a larger proportion of minorities than is typically seen with traditional survey methods, without employing an oversampling strategy [19,20]. This is an especially important strength for research institutions that have limited access to diverse populations in their region. Inadequately representative sampling could be avoided by implementing an oversampling strategy within Facebook's advertising system, partially blocking majority (eg, white, male, younger than 25 years) respondents during data collection.

Most participants reported that they were willing to use Facebook as a therapeutic tool. Most endorsed widespread use of technology, both in terms of devices and services. Participants also reported on a wide range of triggers to their auditory hallucinations, coping skills, and advice to others, providing qualitative insights that can be integrated into the design of technology-based treatment or support systems.

Our results were presented in 2 ways: the phased optimization of advertisement materials and survey design and the quantitative and qualitative results of the survey. First, we

demonstrated that social media outreach strategies can be fine-tuned with systematic testing. Using multiple-image, "carousel"-style advertisements are suggested over single-image advertisements, and it is worthwhile to test a set of images for relative performance. Furthermore, advertisement targeting and restricting open-ended questions to the end of the survey had a positive effect, whereas survey pagination did not matter.

It is important to note that although correlated, that a low cost per click does not directly translate into high data quality. Engagement is necessary but insufficient to collect data from social media. A particular advertisement may generate a large number of clicks, but it not may net as many completed surveys as an advertisement or survey design that performs less well in terms of engagement, but encourages conscientious participation. It is also important to note that comparisons across advertising runs performed at different times are potentially inaccurate due to random fluctuations in advertising performance that are either random or caused by yet unexplored phenomena. Although general differences of large magnitude (eg, a cost per click of $0.10 vs $15.00) will probably hold at different times, smaller differences necessitate the need for comparisons across identical time frames. Comparisons across phases require large differences to be trustworthy, whereas comparisons within phases can be trusted even when the differences are much smaller.

A methodologic shortcoming did not allow us to obtain cost per participant within optimization phases where the factor being manipulated was on the side of Facebook's advertising system, rather than the survey system, as seen in Table 6. Competing advertisements designs registered different tracking metrics within the advertising system, but a single survey was used to collect these data, which washed out the relationship between survey response rates and advertisement click rates. When surveys were compared, separate surveys were used, making these comparisons possible. Future research should consider using separate surveys (or "collectors" in Survey Monkey) to ensure that these data are captured.

Although the use of social media for recruitment offers tremendous new opportunities, the approach has noteworthy limitations. Differential use of social media is the foremost concern between those who have mental health problems and those that do not. More work needs to be done to explore the possible differences in social media use of this population as compared with those who are not experiencing symptoms. Although we know this group reports high social media use, their interaction patterns, use frequency, and posted content may be quite different. Future work should explore any potential differences.

Attention also needs to be paid to the fundamental motivation behind responding to a social media–based advertisement to participate in research. Whatever intrinsic or extrinsic motivational differences differentiate between those that click the advertisement and those that do not may be systematically related to social, psychological, emotional, behavioral, cultural, and demographic factors that impact the ultimate conclusions drawn from research performed with digital recruitment strategies. Regarding our study in particular, we only focused on the images in advertisements, and did nothing to modify the language used in the 25-character title and 90-character description allowed by Facebook. Future studies, complemented by empirically supported best practices from the fields of marketing and advertising, should test advertisement images and wording combinations in a multivariate fashion. Facebook's advertising platform is well suited for such systematic testing with real-time tracking capabilities.

The rapid proliferation of social media has outpaced social and clinical scientists' understanding of how these new resources can be harnessed to improve research, outreach, and services. Services like Facebook, Twitter, Snapchat, and Instagram offer an unprecedented way to reach populations of interest. These platforms automatically record an abundance of social data that can drive new research programs and inform the design of the next generation interventions. The development of best practices and new technology for massive outreach via social media and the collection of social media data itself should be a key focus of current research. This work suggests that engagement with social media advertising is a viable strategy for those that experience auditory hallucinations, hinting at the approach's promise for those that experience other mental health symptoms that may be missed by traditional outreach systems.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Survey items.

[PDF File (Adobe PDF File), 43KB - jmir_v18i6e160_app1.pdf]

References

XSL•FO

RenderX

1. Tien A. Distributions of hallucinations in the population. Soc Psychiatry Psychiatr Epidemiol 1991 Dec;26(6):287-292. [Medline: 1792560]

2. Kessler R, Berglund P, Bruce M, Koch J, Laska E, Leaf P, et al. The prevalence and correlates of untreated serious mental illness. Health Serv Res 2001 Dec;36(6 Pt 1):987-1007 [FREE Full text] [Medline: 11775672]

3. Srinivasa MR, Kishore Kumar KV, Chisholm D, Thomas T, Sekar K, Chandrashekari C. Community outreach for untreated schizophrenia in rural India: a follow-up study of symptoms, disability, family burden and costs. Psychol Med 2005 Mar;35(3):341-351. [Medline: 15841870]

4. Pew Research Center. 2013 Dec 27. Social Networking Fact Sheet URL: http://www.pewinternet.org/fact-sheets/social-networking-fact-sheet/ [accessed 2016-06-01] [WebCite Cache ID 6hwH8IwMc]

5. Snowdy CE, Shoemaker EZ, Chan S, Hilty DM. Social Media and Clinical Practice: What Stays the Same, What Changes, and How to Plan Ahead? In: Mucic D, Hilty DM, editors. e-Mental Health. Switzerland: Springer International Publishing; 2016:151-170.

6. Facebook. 2015. Facebook Company Info URL: https://www.facebook.com/facebook/info/ [accessed 2016-06-01] [WebCite Cache ID 6hwHZRSB8]

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

8. Kapp J, Peters C, Oliver D. Research Recruitment Using Facebook Advertising: Big Potential, Big Challenges. Journal of Cancer Education 2013;28(1):134-137. [Medline: 23292877]

9. Gowen K, Deschaine M, Gruttadara D, Markey D. Young adults with mental health conditions and social networking websites: seeking tools to build community. Psychiatr Rehabil J 2012;35(3):245-250. [doi: 10.2975/35.3.2012.245.250] [Medline: 22246123]

10. Ben-Zeev D, Davis KE, Kaiser S, Krzsos I, Drake RE. Mobile technologies among people with serious mental illness: opportunities for future services. Adm Policy Ment Health 2013 Jul;40(4):340-343 [FREE Full text] [doi: 10.1007/s10488-012-0424-x] [Medline: 22648635]

11. Pedersen E, Helmuth E, Marshall G, Schell T, PunKay M, Kurz J. Using facebook to recruit young adult veterans: online mental health research. JMIR Res Protoc 2015;4(2):e63 [FREE Full text] [doi: 10.2196/resprot.3996] [Medline: 26033209]

12. Kosinski M, Matz S, Gosling S, Popov V, Stillwell D. Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines. Am Psychol 2015 Sep;70(6):543-556. [doi: 10.1037/a0039210] [Medline: 26348336]

13. Fan, Yan. Factors affecting response rates of the web survey: A systematic review. Computers in Human Behavior 2010 Mar;26(2):132-139. [doi: 10.1016/j.chb.2009.10.015]

14. Woodside AG. Measuring Advertising Effectiveness In Destination Marketing Strategies. Journal of Travel Research 1990 Jan 01;29(2):3-8. [doi: 10.1177/004728759002900201]

15. Hsieh H, Shannon S. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]

16. Cassell C, Symon G. Using templates in the thematic analysis of texts. In: Essential guide to qualitative methods in organizational research. London: SAGE Publications; 2004.

17. Ryan GW, Bernard HR. Techniques to identify themes. Field Methods 2003:85-109. [doi: 10.1177/1525822X02239569]

18. Strauss A, Corbin J. Strategies for Qualitative Data Analysis. In: Basics of qualitative research: grounded theory procedures and techniques. Newbury Park, California: Sage Publications; 1990.

19. Bonevski B, Randell M, Paul C, Chapman K, Twyman L, Bryant J, et al. Reaching the hard-to-reach: a systematic review of strategies for improving health and medical research with socially disadvantaged groups. BMC Med Res Methodol 2014;14:42 [FREE Full text] [doi: 10.1186/1471-2288-14-42] [Medline: 24669751]

20. McGraw S, McKinlay J, Crawford S, Costa L, Cohen D. Health survey methods with minority populations: some lessons from recent experience. Ethn Dis 1992;2(3):273-287. [Medline: 1467762]

Abbreviations

SD: standard deviation
SMI: serious mental illness

XSL•FO

RenderX

<u>Original Paper</u>

A Systematic Review of Methods and Procedures Used in Ecological Momentary Assessments of Diet and Physical Activity Research in Youth: An Adapted STROBE Checklist for Reporting EMA Studies (CREMAS)

Yue Liao[1*], PhD, MPH; Kara Skelton[2*], MA; Genevieve Dunton[3*], PhD; Meg Bruening[4*], PhD, MPH, RD

[1]University of Texas MD Anderson Cancer Center, Department of Behavioral Science, Houston, TX, United States

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

[3]University of Southern California, Department of Preventive Medicine, Los Angeles, CA, United States

[4]Arizona State University, School of Nutrition and Health Promotion, Phoenix, AZ, United States

[*]all authors contributed equally

Corresponding Author:
Meg Bruening, PhD, MPH, RD
Arizona State University
School of Nutrition and Health Promotion
500 N 5th Street
Phoenix, AZ, 85004
United States
Phone: 1 6028272266
Fax: 1 602.8272253
Email: Meg.Bruening@asu.edu

Abstract

Background: Ecological momentary assessment (EMA) is a method of collecting real-time data based on careful timing, repeated measures, and observations that take place in a participant's typical environment. Due to methodological advantages and rapid advancement in mobile technologies in recent years, more studies have adopted EMA in addressing topics of nutrition and physical activity in youth.

Objective: The aim of this systematic review is to describe EMA methodology that has been used in studies addressing nutrition and physical activity in youth and provide a comprehensive checklist for reporting EMA studies.

Methods: Thirteen studies were reviewed and analyzed for the following 5 areas of EMA methodology: (1) sampling and measures, (2) schedule, (3) technology and administration, (4) prompting strategy, and (5) response and compliance.

Results: Results of this review showed a wide variability in the design and reporting of EMA studies in nutrition and physical activity among youth. The majority of studies (69%) monitored their participants during one period of time, although the monitoring period ranged from 4 to 14 days, and EMA surveys ranged from 2 to 68 times per day. More than half (54%) of the studies employed some type of electronic technology. Most (85%) of the studies used interval-contingent prompting strategy. For studies that utilized electronic devices with interval-contingent prompting strategy, none reported the actual number of EMA prompts received by participants out of the intended number of prompts. About half (46%) of the studies failed to report information about EMA compliance rates. For those who reported, compliance rates ranged from 44-96%, with an average of 71%.

Conclusions: Findings from this review suggest that in order to identify best practices for EMA methodology in nutrition and physical activity research among youth, more standardized EMA reporting is needed. Missing the key information about EMA design features and participant compliance might lead to misinterpretation of results. Future nutrition and physical activity EMA studies need to be more rigorous and thorough in descriptions of methodology and results. A reporting checklist was developed with the goal of enhancing reliability, efficacy, and overall interpretation of the findings for future studies that use EMAs.

Introduction

The number of overweight or obese youth in the United States is alarming for public health professionals, as prevalence of overweight/obesity among youth is estimated to be 31.8% [1]. National data suggests that only 15.7% of adolescents ate vegetables 3 or more times during the past 7 days and only 29% of adolescents achieved 60 minutes of physical activity per day [2]. US children and adolescents' lifestyle factors, such as poor diet and physical inactivity, are related to an increased risk for chronic diseases, including diabetes, hypertension, cardiovascular diseases, and other metabolic disorders [3,4]. Many current methods for assessing nutrition and physical activity (eg, dietary recalls, physical activity logs) are limited since they can introduce high participant burden [5] and are prone to inaccuracies. More studies that use assessment methods that may limit participant burden and provide ecologically valid data for nutrition and physical activity behaviors are needed.

Advances in electronic technologies and societal changes have created opportunities to assess youth nutrition and physical activity behaviors as they occur in their daily lives. Real-time data capture methods refer to collecting data as it naturally occurs [6]. Real-time data assessments differ from traditional retrospective data collection methods as they sample snapshots of participants' lives to capture the variability of experiences more accurately. As information is collected at or near the moment when events and experiences occur, real-time data capture methods can reduce memory and other biases that are associated with retrospective recall measures [7]. Ecological momentary assessment (EMA), a type of real-time data capture method, was originally developed for psychological assessments of mood and affect [8]. Shiffman and colleagues [5] define EMA as "monitoring or sampling strategies to assess phenomena at that moment they occur in natural settings." There are several unique features common to the EMA methods: (1) the data capture happens in subjects' natural environment—the "Ecological" aspect of EMAs; (2) assessments focus on current feelings and behaviors, rather than concentrating on recall or summary over long periods of time—the "Momentary" aspect of EMAs; (3) the moments are assessed by random sampling, event-based sampling, interval sampling, or a combination of any of these strategies; and (4) multiple assessments are collected over time to provide a profile for behavior throughout time—the "Assessment" aspect of EMAs [5].

Nutrition and physical activity studies that employ the EMA methodology enable the collection of data with an array of variables including behavioral, physical, sociopsychological, and contextual information [8]. This assessment strategy makes it possible to examine concurrent exposures and events, such as examining where and with whom physical activity and sedentary behavior are likely to occur during the course of participants' everyday lives [9]. Due to the repeated measurements used in EMA methodology, EMA studies are able to focus on within-person changes in behaviors and experiences over time, thus allowing the investigation of antecedents and consequences of a behavior [10], and the advanced modeling of how variation in momentary cognitive state might relate to behaviors [11].

Over the past several years, there has been an increase in the popularity and prevalence of research conducted using EMAs. Given the potential methodological and analytical advantages of using EMAs in nutrition and physical activity research in youth, this review is aimed to describe features of EMA methodology in studies that address nutrition and physical activity in children and adolescents. In addition, although some guidance is available for designing and reporting in EMA studies [12], there are currently no specific guidelines for the necessary detail in reporting in EMA studies, which could make a systematic synthesis of results from EMA studies challenging. Similar reporting checklists for other types of studies have been widely adopted. For example, the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) is a commonly used checklist of items for observational studies [13]. It contains 22 items that relate to the title, abstract, introduction, methods, results, and discussion sections of papers with the goal to improve the quality of reporting. Building on the STROBE checklist and the EMA design guidelines by Stone and Shiffman [12], a comprehensive checklist of specific items to be reported for EMA studies was also developed: Checklist for Reporting EMA Studies (CREMAS).

Methods

Information Sources

CINAHL, PsychINFO, PubMed, and EBSCOhost were searched for relevant studies that were published before July 2015. The keywords used included "ecological momentary assessment" and "EMA," in combination with "food," "nutrition," "eating," "food consumption," "eating habits," "physical activity; PA," "text messaging," "SMS telephone," "electronic diaries," and "prompting." A hand search of the reference section of all papers was conducted to review for additional papers that were missed during the electronic search.

Selection Criteria

Inclusion criteria for this review were as follows: (1) published in English, (2) used EMA-based data collection method, (3) had a mean participant age of 22 years or younger (or enrolled in a college/university), and (4) focused on the assessment of nutrition or physical activity habits. Studies were excluded if they did not have repeated measures, did not assess variable/outcome measures via EMAs, had a mean participant age greater than 22 years, assessed maladaptive or disordered nutrition or physical activity behaviors, and/or were intervention studies. Further, papers must have reported results of EMAs; papers that only described EMA design were not included.

Data Extraction

Data were extracted in two passes. In the first pass, data pertaining to the following general study characteristics were extracted from each of the studies: sample size, study design,

measures, research questions/objectives, findings, and limitations/future directions. In the second pass, data extraction continued by gathering specific methodological features and response- and compliance-related information. In particular, data were synthesized from the following 5 main areas:

1. *Sampling and measures*: sample characteristics and tools used in the EMA protocol

2. *Schedule*: monitoring periods (number of waves from which data were collected), duration (number of days that each monitoring period lasted), prompt frequency (frequency of EMA prompts per day), and prompt interval (the time between each EMA prompt)

3. *Technology and administration*: use or lack of technology and method of administration of EMAs

4. *Prompting strategy*: methods used to cue participants—interval contingent (EMA prompts were set for certain intervals that were not random), random interval contingent (EMA prompts were set to be randomized throughout the day), event based (EMAs were recorded when eating occasions or physical activity occurred), or evening report (EMAs administered in the evenings to summarize the events of the day)

5. *Response and compliance*: participation rate, gathered data, missing data (ie, unanswered and/or unprompted EMA surveys), latency (ie, the time period between when participants receive an EMA prompt and when the EMA is answered), and attrition (ie, the number of participants who dropped out of the study for any reason).

For studies that did not report any of this data, calculations were performed using information provided in the paper whenever possible.

A coding form was developed based on the above areas of interest and two raters extracted information from each study independently for all items. Agreement among raters for each item ranged from 85 to 100%, and all discrepancies were resolved through discussions that led to consensus.

Results

Literature Search

After completing a systematic review of databases and reference lists, a total of 428 potentially relevant studies were screened. From this group, 62 article abstracts were identified and evaluated for inclusion criteria, and 23 were selected for further full-text review. In cases where multiple papers were published (*n*=5) on the same study (eg, reporting validity, reporting outcomes, translations, etc.), information was extracted from all papers and presented as a single study. On the basis of the abovementioned criteria, 13 independent papers were retained for inclusion in the review, 7 studies were physical activity-related, 5 focused on nutrition outcomes, and 1 study assessed both physical activity and nutrition behaviors. Figure 1 presents a flow chart of the systematic literature search, according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.

Sampling and Measures

Sample characteristics and methodological features of each included study are presented in Table 1. The mean number of participants per study was 391 (range=30-1604, median=147). However, this mean is skewed by 5 studies with samples over 500 [14-18]. Excluding these 5 studies, the mean sample size was 82 (range=30-158, median=63). Two studies did not report mean age [19,20]. Excluding those 2 studies, the mean sample-weighted age of participants was 15.6 years, with a range of 5.3-21.0 years.

One study asked participants to respond to the question, "What are you doing now?" All other studies used retrospective questions (ranging from every 15 minutes to 4 hours) to assess nutrition and physical activity behaviors. Only one study combined EMA with an objective measurement (ie, accelerometry) [21].

Figure 1. PRISMA Flow Diagram for paper selection process.
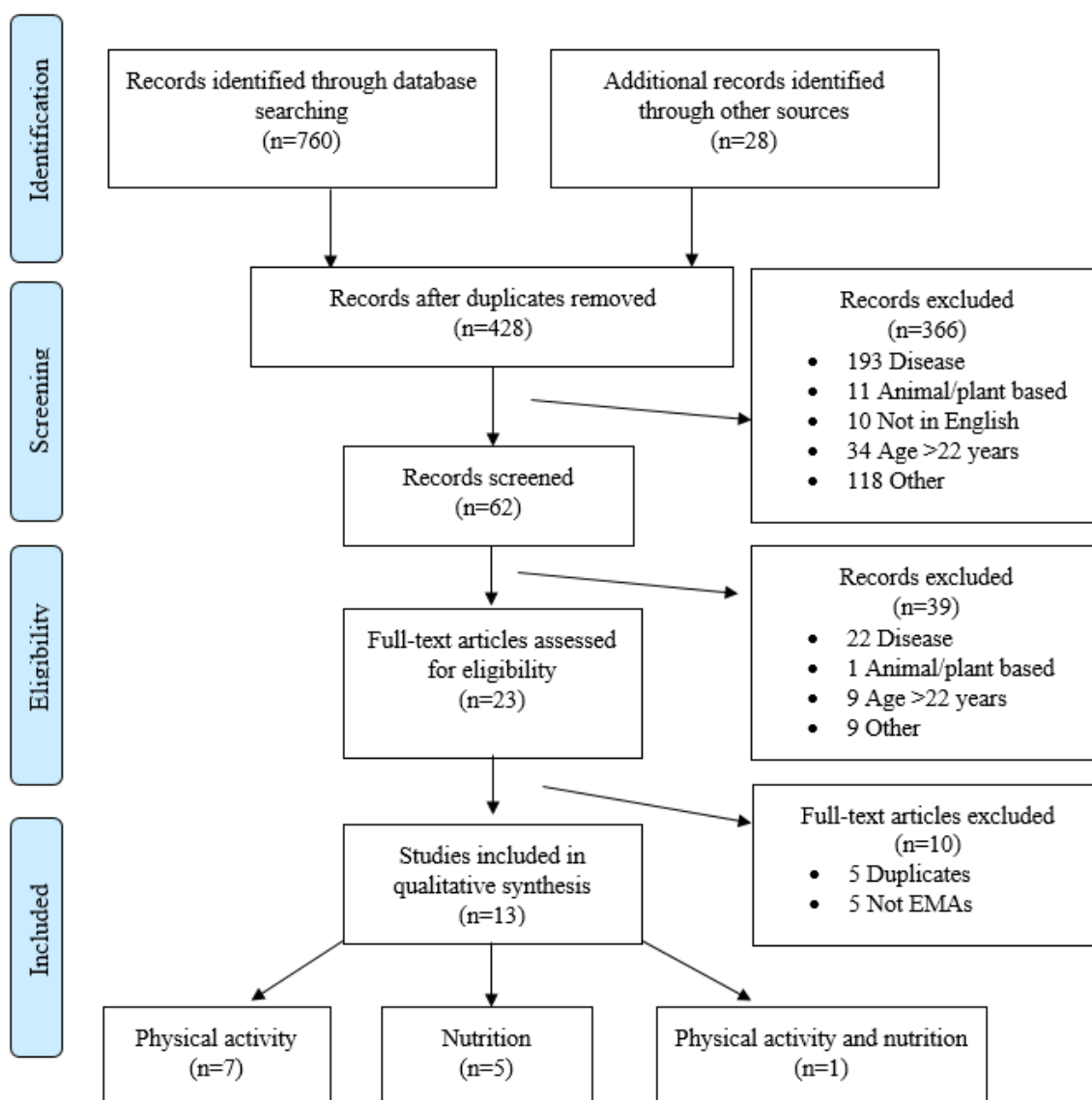
Table 1. Methodological features of ecological momentary assessment (EMA) nutrition and physical activity studies in youth.

Citation	Technology[a]	Prompt approach[b]	Monitoring periods[c]	Duration (days) per monitoring period[d]	Prompt frequency per day[e]	Prompt interval[f]
Berkman et al [22]	Paper-and-pencil diary and cell phone	Event-based	1	14	4	Breakfast, lunch, dinner, and bedtime predefined by participants
Biddle et al [14]	Paper-and-pencil diary	Fixed interval contingent	1	4	44 weekdays 68 weekends	15 minutes
Biddle et al [15]	Paper-and-pencil diary	Interval contingent	1	4	44 weekdays 68 weekends	15 minutes
Carels et al [23]	Paper-and-pencil diary	Event-based & random interval contingent	1	7	-	<15 minutes of event
Dunton et al [16]	Palm III handheld computer	Fixed interval contingent	8	4	20-30	30 minutes (+ 10 minutes)
Dunton et al [24]	HTC Shadow cell phone	Random interval contingent	2	4	3 weekdays 7 weekends	Random within 2-hour blocks
Gorely et al [17]	Paper-and-pencil diary	Fixed interval contingent	2	4	44 weekdays 68 weekends	15 minutes
Grenard et al [25]	Palm E2 PCA handheld computer	Event-based, fixed interval contingent, and evening report	1	7	2 weekday 4 weekends	Event-based: < 15 minutes of event Fixed interval: 3 hours
Mak et al [18]	Paper-and-pencil diary	Event-based	1	4	7	3 hours
Rouse et al [26]	Paper-and-pencil diary	Fixed interval contingent	1	2	44 weekdays 68 weekends	15 minutes
Rusby et al [19]	iPod touch handheld computer	Random interval contingent	4	7	3 M-T 4 F 6 Sat. 5 Sun	90-120 minutes
Spook et al [20]	Blackberry OS, Android, iOS, mobile phones	Event-based and interval contingent	1	7	5	3-4 hours
Thomas et al [27]	Palm-top handheld computer	Random interval contingent	1	7	6	Variable

[a]Technology: operating system, device type, and/or phone model (in as much detail as was provided in the paper).

[b]Prompt approach: type of EMA sampling.

[c]Monitoring periods: number of waves EMA was used in the study.

[d]Duration: number of days each monitoring period lasted.

[e]Prompt frequency: number of times it was intended for participants to answer EMA prompts.

[f]Prompt interval: time between each EMA prompt.

Schedule

The majority of studies (9 out of 13) monitored participants during one period of time (ie, one wave of data collection), while other studies included up to 8 waves of data collection. The duration of each monitoring period included: 2 days (1 study), 4 days (6 studies), 7 days (5 studies), and 14 days (1 study). Studies with more than one monitoring period had smaller durations than those with only one monitoring period. Typically, the shorter the duration of the study, the higher the prompt frequency per day. For example, one study prompted participants 44 times on weekdays and 68 times on weekends during a 4-day monitoring period [17]. The studies with the longest durations prompted their participants 4 times per day for 14 consecutive days [15].

Seven of the reviewed studies had different prompting frequencies for weekdays and weekend days. In general, participants received more prompts during weekend days than weekdays. Prompting frequency ranged from 2 times per day (during weekdays) to 68 times per day (during weekend days),

XSL•FO

RenderX

with the median being 7 times per day. The majority of the studies (9 out of 13) did not collect EMA data during school hours (eg, between 8am-3pm). There were several studies conducted by the same group of researchers that used the same prompting frequency schedule across studies: 44 prompts per day during weekdays and 68 prompts during weekend days [14,15,26,28]. Prompt frequencies varied significantly between studies that employed paper and pencil as compared with electronic data collection tools. For example, a paper-and-pencil study [18] utilized a prompt frequency of 7 times per day, while an electronic data collection EMA study prompted participants 20-30 times per day [16].

Technology and Administration

A majority of the studies (7 out of 13) used electronic EMA methods and the rest of the studies used paper-and-pencil-based diary methods. For studies that used electronic EMA methods, 4 used cellular phones and 3 used handheld computers. Only one study used a combination of technology for the EMAs. This study divided the sample into two equal groups: one group completed EMAs via paper-based diary and the other group completed EMAs via cellular phones [22]. With the exception of one study [16], studies that used electronic devices for the EMAs had relatively small sample sizes (n<175). Training sessions on the use of the EMA technology for participants were held in two studies. One study used parent-reported dietary consumption data for children aged 1.5-10 years old [18]; all other studies collected the self-reported data directly from the youth.

Prompting Strategy

Most (11 out of 13) of the studies used interval-contingent prompting strategy. Of these 11 studies, 5 used fixed interval contingent only (eg, every 15 minutes), 3 used random interval contingent only (eg, randomly within a 2-hour block), and 3 used combined strategies (eg, event based and interval contingent, interval contingent and evening report). One study used event-based strategy only for collecting EMA responses.

The sampling strategy used by the studies seemed to be related to the behavior of interest. All but one study that measured physical activity used interval contingent sampling, while the

majority of studies measuring nutrition habits used events-based sampling in their design.

Response and Compliance

Table 2 summarizes the response and compliance-related results for all studies. Although most studies reported participant initial enrollment, only 2 studies formally reported attrition rate [19,20]. Another 9 studies reported their respective analytical sample size, although most of the studies did not clearly indicate why the analytical sample size varied from the initial enrollment (eg, participant attrition, device malfunction, or other reasons).

For the studies that utilized an interval-contingent prompting strategy via electronic devices, none of the studies reported how many prompts were actually received by participants. Eight studies did not report the average number of percentage of EMA prompts answered by participants. No study reported reasons for unprompted or unanswered prompts.

Among studies that reported compliance, compliance rates were relatively high (mean=71.3%), with reported compliance ranging from 43.8-95.9%. Compliance reporting differed for paper-and-pencil and electronic EMA designs. Only 2 (out of 6) paper-and-pencil designs reported compliance [26], whereas all of the electronic designs reported compliance rates. Results from Berkman et al compared compliance between paper-and-pencil and electronic EMAs, and reported that the electronic group was more compliant than the paper-and-pencil group (95.9 and 69.9%, respectively) [22]. Even though several studies had more than one monitoring period, no studies reported compliance by wave. One study reported compliance by day [20] and reported that daily average compliance rates declined from 63% at the start of the study to 23% on day 7, demonstrating a decline in answered EMA prompts as the monitoring period progressed.

Only 3 studies reported latency (the time period between when participants receive an EMA prompt and when the EMA is answered) of participant responses. In order to ensure the momentary nature of the responses, 2 electronically administered EMA studies designed their EMAs to prohibit responses 4 minutes [16] or 8 minutes [19] after signaling prompts were sent. No studies reported on why respondents were late responding to prompts.

Table 2. Ecological momentary assessment (EMA) response and compliance-related results from nutrition and physical activity studies in youth.

Citation	Initial enrollment[a]	Analytical sample size[b]	Average answered EMA survey prompts (per participant) M (SD)[c]	Average compliance rate[e]	Average latency (>15 minutes)[f]
Berkman et al [22]	44	NR[fg]	NR[fg]	Electronic: 96% Paper and pencil: 70%	Electronic: 40.1% Paper and pencil: 73.2%
Biddle et al [14]	991	948	NR[fg]	NR[fg]	71.7%
Biddle et al [15]	623	550	NR[fg]	NR[fg]	NR[fg]
Carels et al [23]	30	NR[fg]	Lapses: 11.8 (10.9) Temptations: 8.7 (8.3) Random prompts: 18.3 (8.3)	NR[fg]	N/A
Dunton et al [16]	568	524	24.3 (3.4)	83% (SD=9.4)	0%
Dunton et al [24]	121	108	31.2[a]	78%	NR[fg]
Gorely et al [17]	1604	1371	NR[fg]	50%	74.1%
Grenard et al [25]	158	158	Random: 11.8[a] Eating events: 13.4[a] Evening report: 6.58[a]	Random: 71% Evening reports: 95%	NR[fg]
Mak et al [18]	-	642	N/A	NR[fg]	N/A
Rouse et al [26]	147	84	NR[fg]	57%	NR[fg]
Rusby et al [19]	82	80	74.9[a]	Total: 69%[a]	0%
Spook et al [20]	30	30	4.3	44%	NR[fg]
Thomas et al [27]	43	39	31.3%[a]	71%	NR[fg]

[a]Initial enrollment: number of participants who consented to the study.

[b]Analytical sample size: number of participants in the main analysis.

[c]Average answered EMA survey prompts (per participant): average of number of survey prompts each participant responded to.

[d]Average compliance rate: average of number of answered surveys out of total planned EMA surveys per participant, can include compliance for each monitoring period.

[e]Average latency (>15 minutes): the average time between prompting to participants answered the prompt.

[f]Numbers were hand calculated from information available.

[g]NR[g]: not reported in paper.

Discussion

The primary aim of this study was to systematically review the literature on EMA methods and procedures relating to nutrition and physical activity in youth in order to describe the common practices in EMA methodologies, and to identify response and compliance rates for this target population group. There has been very limited research using EMA methodology to assess youth nutrition and physical activity behaviors. A total of 13 individual EMA studies met inclusion criteria for this review and varied considerably in methodological and results reporting strategy. Enhancements to design and reporting may increase the interpretability and generalizability of EMA findings, application to intervention projects, and ease of use when assessing nutrition and physical activity among youth.

Overall, a significant amount of key information was not reported from studies that were included in this review, demonstrating the need for a reporting guideline that is tailored to the unique features of EMA studies, especially studies that utilize electronic devices. On the basis of results from this review and building on existing guidelines [13], CREMAS was developed to provide recommendations in reporting future EMA studies (Table 3). These recommendations to unify reporting include 16 items that address various sections in a manuscript, and in general, could be applied to EMA studies across disciplines.

Table 3. An adapted STROBE Checklist for Reporting EMA Studies (CREMAS).

Topic	Item #	Checklist item	Page number reported
Title			
	1		Include ecological momentary assessment in title and key words
Introduction			
Rationale	2		Briefly introduce the concept of EMA and provide reasons for utilizing EMA for this study or topic of interests (eg, to examine time-varying predictors of unhealthy eating occasions in children's daily lives)
Methods[a]			
Training	3		Indicate if, and by what methods, training of participants for EMA protocol was used
Technology	4		Describe what technology, if any, was used. Include the following information: device (eg, mobile phone, portable computer), model (eg, Nexus 4, iPod), operating system (eg, Android, Windows), and EMA program name
Wave duration	5		State the number of waves for the study (eg, 2 monitoring periods over the course of 1 year)
Monitoring period	6		State the number of days each wave of the study lasted, and how many weekdays versus weekend days
Prompting design	7		Indicate the prompting strategy used for the study (eg, event-based, interval-based, or a combination of the two). If using interval-based strategy, indicate what type of schedule is used (eg, fixed, random, or hybrid interval)
Prompt frequency	8		Intended frequency of prompts per day. Break down by weekdays and weekend days if applicable
Design features	9		Describe any design feature to address potential sources of bias (eg, reactivity) or participant burden (eg, EMA questions appearing in different orders)
Results[a]			
Attrition	10		Indicate participant attrition throughout the study; report attrition rates both by monitoring days and waves, if applicable
Prompt delivery	11		Report number of EMA prompts that were planned to be delivered. If possible, also report the number of EMA prompts that were actually received by participants and indicate reasons for why prompts were not sent out (eg, technical issues or participant noncompliance reason such as phone was powered off)
Latency	12		Report the amount of time from prompt signal to answering of prompt
Compliance rate	13		Report total answered EMA prompts across all subjects and the average number of EMA prompts answered per person. Report compliance rate both by monitoring days and waves, if applicable. Indicate reasons for noncompliance, if known
Missing data	14		Report whether EMA compliance is related to demographic or time-varying variables
Discussion			
Limitations	15		Discuss limitations of the study, taking into account sources of potential bias when using EMA methods (eg, reactivity, use of technology)
Conclusions	16		Provide a general interpretation of results and discuss the benefits of using EMA (eg, improving understanding of daily behaviors)

[a]Secondary data analysis paper can refer to a main methods paper that has discussed all of these items.

This review shows that studies have used both paper-and-pencil and electronic EMA designs to capture nutrition and physical activity in youth. Compared with paper-and-pencil design, there are several benefits to using electronic EMA designs such as automatic prompt signaling (eg, auditory or tactile), instant data transfer via download or Internet-based secure servers, and greater accessibility and convenience for participants [29]. More importantly, electronic EMA collection instruments are able to collect exact times of each assessment and ensure that the assessments are completed following study protocols. Indeed, several studies have shown that even with signaling prompts and detailed instructions, the completion of paper-and-pencil EMAs may not occur in real time [30].

Although technologies can assist in making the delivery of EMA surveys more systematic, they also have some limitations. For example, EMA surveys may fail to be delivered because of technological issues (eg, problems with the app) or user

XSL•FO
RenderX

compliance issues (eg, subjects can have phone turned off). Therefore, it is important that authors report the intended number of EMA prompts and the actual number of EMA prompts participants received if possible. The utilization of electronic EMA devices may pose some challenges for some studies. For example, the electronic devices themselves can be a costly research expense. The majority of the reviewed studies that used electronic devices provided those devices to the participants (instead of participants using their own device), which ensured the consistency of the usability and functionality in the administration of EMAs. However, given the cost associated with providing loaned devices, drawbacks, such as limiting sample size and participant burden of remembering to keep the study device with them and charged, should be considered. Often, an experienced computer programmer and several rounds of pilot testing are needed to develop electronic EMAs to be administered on mobile phones or personal digital assistants (PDAs). Nevertheless, free open source EMA programs (eg, PACO by Google, MovisensXS by Movisens GmbH) are available and can be tailored to researchers' specifications.

In general, response and compliance-related data were inconsistently reported. This information is critical to assess the quality of data collected by a study. More importantly, these data will provide valuable information for future studies planning to adopt EMA methods in optimizing study design (eg, Will compliance rate be very different between a study that delivers 4 prompts a day versus a study that delivers 44 prompts a day?). This review is not able to answer this question fully because over half of the reviewed studies did not report compliance data. Therefore, it is highly recommended that all future EMA studies report response and compliance-related data, except for studies that only utilize event-based design with manually initiated reporting since there is no set number of diary entries or prompts that participants are required to complete. In addition, the majority of studies did not report latency. Due to the in-the-moment nature of EMA studies, it is critical that EMAs are completed shortly after prompts are received. One way to ensure the momentary nature of the responses is limit the time respondents have to complete the EMA, as was done by two studies included in this review [16,19].

Although there is no consistently agreed upon gold standard for acceptable rates of compliance to EMAs, Stone and Shiffman noted that if compliance falls below 80% there may be concern that data are not representative or generalizable to participants' usual daily lives [12]; however, reasons for missing data (random vs. not random) should be taken into consideration. Thus, we encourage future EMA studies to report reasons for noncompliance or missing data whenever possible. Regardless of compliance rate, missing EMA data (prompted and unprompted) should be examined for systematic associations with known temporal (eg, time of day, day of the week, chronological day in study, study wave) and demographic (eg, age, gender, race/ethnicity, SES, adiposity) factors [31,32]. A more thorough analysis of missing EMA data would include examining whether the rates and likelihood of unanswered EMA prompts are associated with information provided by temporally adjacent available EMA data (eg, average daily levels, levels

reported at EMA prompts before or after the unanswered prompt). Pattern-mixture random-effects regression modeling offers a promising strategy for understanding missingness patterns with EMA data [33]. For data determined to be missing at random (MAR) or missing completely at random (MCAR) (ie, associated with unobserved or observed variables), imputation methods should be considered [34]. With consistent reporting of response and compliance rates, audiences would be able to determine whether the data may be generalizable to all days of the week, times of day, or situations throughout the day.

Even though EMAs offer many methodological benefits, there are still some challenges when utilizing real-time data capture methods. Although most EMA studies aim to observe participants' behavior without influencing it, repetitive exposure to EMA items relating to nutrition and physical activity may trigger participants to adjust behaviors in ways they otherwise would not. Evidence suggests that the mere act of measuring a behavior could have some impact on that behavior in the future [35]. Further, if EMA prompting rates are too frequent and/or EMA questions are too repetitive, participants may opt not to respond to the surveys or drop out of the study altogether. The study with the highest frequency of prompting (44 prompts during weekdays and 68 prompts during weekend days) also reported the lowest compliance rate at 57% [26]. To reduce concerns about participant reactivity and burden, researchers should aim to use the fewest number of prompted surveys possible to answer their questions or interests.

Researchers could also consider combining EMAs with other objective measurement to capture the behaviors of interests. For example, Dunton and colleagues used electronic EMA in combination with accelerometry to measure children's physical activity [21]. In this case, the accelerometry device can continuously measure activity intensity while EMAs can be used to capture other information such as type of activity, and contextual information of activities (eg, where and with whom).

Overall, the lack of consistency in reporting EMA methods greatly limits the scientific impact and possible use of findings for behavioral assessments or development of intervention strategies for nutrition and physical activity behaviors in youth. A clear and detailed report of EMA design features could be very helpful for researchers that are new to EMA methodologies. Consistently reporting these types of data will also be useful for future researchers to understand which device/model/systems are effective for nutrition and physical activity assessment studies. Without providing important aspects of EMA design and results, data can be misinterpreted. Researchers may also want to report intrapersonal (person-level) compliance rates, as there might be significant individual variation. In general, reporting more complete aspects of EMA data will help the audience to fully interpret the results, including generalizability and application to future EMA designs.

Limitations

Although this review is unique in that it is the first to examine EMA studies of nutrition and physical activity behaviors among youth, it has several limitations. First, we attempted to be exhaustive in the literature search, but it is possible that some

studies may have been missed. Second, since reporting strategies were so diverse, our ability to report quantitative information was limited. Further, for total EMA prompts received and answered, latency, compliance, and attrition rates, so much data was missing across studies that it was hard to make intuitive interpretations of these results.

Conclusions

This review presented the data of key EMA methods from 13 nutrition and physical activity studies. Utilizing EMA methods to study nutrition and physical activity in young people has many powerful benefits, including ecological validity and minimizing retrospective response bias. However, based on our review, many studies fail to employ all the features of EMA methods, as described by Shiffman and colleagues [5], and reporting strategies are inconsistent and insufficient. In order to maximize the impact that EMA data has in the scientific literature, reporting needs to be systematic across studies, allowing greater interpretability and reach of EMA methodologies. Therefore, in order to adequately interpret findings from EMA studies, several items need to be included when reporting EMA methods and results; we created a checklist for others to use. Reporting these key methodological EMA data can enhance efficacy, reliability, and validity of study findings and may lead to increased understanding and interpretation of results.

Conflicts of Interest

None declared.

References

1. Ogden CL, Carroll MD, Curtin LR, Lamb MM, Flegal KM. Prevalence of high body mass index in US children and adolescents, 2007-2008. JAMA 2010 Jan;303(3):242-249. [doi: 10.1001/jama.2009.2012] [Medline: 20071470]
2. Kann L, Kinchen S, Shanklin SL, Flint KH, Hawkins J, Harris W, et al. Center for Surveillance, Epidemiology, Laboratory Services, CDC. 2013 Jun 13. Youth Risk Behavior Surveillance- United States, 2013 URL: http://www.cdc.gov/mmwr/preview/mmwrhtml/ss6304a1.htm [accessed 2016-05-18] [WebCite Cache ID 6hjLFjT0H]
3. Bridger T. Childhood obesity and cardiovascular disease. Paediatr Child Health 2009 Mar;14(3):177-182 [FREE Full text] [Medline: 20190900]
4. Sweat V, Bruzzese J, Albert S, Pinero DJ, Fierman A, Convit A. The Banishing Obesity and Diabetes in Youth (BODY) Project: description and feasibility of a program to halt obesity-associated disease among urban high school students. J Community Health 2012 Apr;37(2):365-371. [doi: 10.1007/s10900-011-9453-8] [Medline: 21826529]
5. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [Medline: 18509902]
6. Stone AA, Broderick J. Real-time data collection for pain: appraisal and current status. Pain Med 2007 Oct;8:S85-S93. [Medline: 17877531]
7. Schwarz N. Retrospective and concurrent self-reports: the rationale for real-time data capture. In: The science of real-time data capture: self-reports in health research. Oxford: Oxford University Press; 2007:11-26.
8. Dunton GF. Using real-time data capture methods to investigate children's physical activity and eating behaviors. In: Segel CM , editor. Childhood obesity: risk factors, health effects and prevention. New York: Nova Science Publishers; 2011:105-114.
9. Liao Y, Intille SS, Dunton GF. Using ecological momentary assessment to understand where and with whom adults' physical and sedentary activity occur. Int J Behav Med 2015 Feb;22(1):51-61. [doi: 10.1007/s12529-014-9400-z] [Medline: 24639067]
10. Liao Y, Shonkoff ET, Dunton GF. The Acute Relationships Between Affect, Physical Feeling States, and Physical Activity in Daily Life: A Review of Current Evidence. Front Psychol 2015;6:1975 [FREE Full text] [doi: 10.3389/fpsyg.2015.01975] [Medline: 26779049]
11. Pickering TA, Huh J, Intille S, Liao Y, Pentz MA, Dunton GF. Physical activity and variation in momentary behavioral cognitions: an ecological momentary assessment study. J Phys Act Health 2016 Mar;13(3):344-351. [Medline: 26284314]
12. Stone AA, Shiffman S. Capturing momentary, self-report data: a proposal for reporting guidelines. Ann Behav Med 2002;24(3):236-243. [Medline: 12173681]
13. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Prev Med 2007 Oct;45(4):247-251. [doi: 10.1016/j.ypmed.2007.08.012] [Medline: 17950122]

XSL•FO

RenderX

14. Biddle SJ, Gorely T, Marshall SJ, Cameron N. The prevalence of sedentary behavior and physical activity in leisure time: a study of Scottish adolescents using ecological momentary assessment. Prev Med 2009 Feb;48(2):151-155. [Medline: 19046984]

15. Biddle SJ, Soos I, Hamar P, Sandor I, Simonek J, Karsai I. Physical activity and sedentary behaviours in youth: data from three Central-Eastern European countries. Eur J Sport Sci 2009 Sep;9(5):295-301. [doi: 10.1080/17461390902829234]

16. Dunton GF, Whalen CK, Jamner LD, Floro JN. Mapping the social and physical contexts of physical activity across adolescence using ecological momentary assessment. Ann Behav Med 2007 Oct;34(2):144-153. [Medline: 17927553]

17. Gorely T, Marshall SJ, Biddle SJ, Cameron N. The prevalence of leisure time sedentary behaviour and physical activity in adolescent girls: an ecological momentary assessment approach. Int J Pediatr Obes 2007 Dec;2(4):227-234. [Medline: 17852544]

18. Mak TN, Prynne CJ, Cole D, Fitt E, Roberts C, Bates B, et al. Assessing eating context and fruit and vegetable consumption in children: new methods using food diaries in the UK National Diet and Nutrition Survey Rolling Programme. Int J Behav Nutr Phys Act 2013;9:126. [doi: 10.1186/1479-5868-9-126]

19. Rusby JC, Westling E, Crowley R, Light JM. Psychosocial correlates of physical and sedentary activities of early adolescent youth. Health Educ Behav 2014 Feb;41(1):42-51. [Medline: 23640122]

20. Spook JE, Paulussen T, Kok G, Van Empelen P. Monitoring dietary intake and physical activity electronically: feasibility, usability, and ecological validity of a mobile-based ecological momentary assessment tool. J Med Internet Res 2013 Sep;15(9):e214. [Medline: 24067298]

21. Dunton GF, Liao Y, Intille SS, Spruijt-Metz D, Pentz M. Investigating children's physical activity and sedentary behavior using ecological momentary assessment with mobile phones. Pediatr Obes 2011 Jun;19(6):1205-1212. [doi: 10.1038/oby2010.302]

22. Berkman ET, Giuliani NR, Pruitt AK. Comparison of text messaging and paper-and-pencil for ecological momentary assessment of food craving and intake. Appetite 2014 Oct;81:131-137 [FREE Full text] [doi: 10.1016/j.appet.2014.06.010] [Medline: 24930596]

23. Carels RA, Hoffman J, Collins A, Raber AC, Cacciapaglia H, O'Brien WH. Ecological momentary assessment of temptation and lapse in dieting. Eat Behav 2001;2(4):307-321. [Medline: 15001025]

24. Dunton GF, Intille SS, Wolch J, Pentz MA. Children's perceptions of physical activity environments captured through ecological momentary assessment: a validation study. Prev Med 2012 Aug;55(2):119-121. [doi: 10.1016/j.ypmed.2012.05.015] [Medline: 22659225]

25. Grenard JL, Stacy AW, Shiffman S, Baraldi AN, MacKinnon DP, Lockhart G, et al. Sweetened drink and snacking cues in adolescents: a study using ecological momentary assessment. Appetite 2013 Aug;67:61-73 [FREE Full text] [doi: 10.1016/j.appet.2013.03.016] [Medline: 23583312]

26. Rouse PC, Biddle S. An ecological momentary assessment of the physical activity and sedentary behaviour patterns of university students. Health Education Journal 2010 Mar 26;69(1):116-125. [doi: 10.1177/0017896910363145]

27. Thomas JG, Doshi S, Crosby RD, Lowe MR. Ecological momentary assessment of obesogenic eating behavior: combining person-specific and environmental predictors. Obesity (Silver Spring) 2011 Aug;19(8):1574-1579 [FREE Full text] [doi: 10.1038/oby.2010.335] [Medline: 21273995]

28. Gorely T, Marshall SJ, Biddle SJ, Cameron N. Patterns of sedentary behaviour and physical activity among adolescents in the United Kingdom: Project STIL. J Behav Med 2007 Dec;30(6):521-531. [doi: 10.1007/s10865-007-9126-3] [Medline: 17712619]

29. Heron KE, Smyth JM. Body Image Discrepancy and Negative Affect in Women's Everyday Lives: An Ecological Momentary Assessment Evaluation of Self-Discrepancy Theory. Journal of Social and Clinical Psychology 2013 Mar;32(3):276-295. [doi: 10.1521/jscp.2013.32.3.276]

30. Broderick JE, Schwartz JE, Shiffman S, Hufford MR, Stone AA. Signaling does not adequately improve diary compliance. Ann Behav Med 2003 Oct;26(2):139-148. [Medline: 14534031]

31. Little R, Rubin D. When questions change behavior: the role of ease of representation. In: Statistical analysis with missing data. Hoboken, NJ: Wiley; 2002.

32. Schafer J. Analysis of incomplete multivariate data. London: Chapman & Hall; 1997.

33. Hedeker D, Rose JS. The natural history of smoking: a pattern-mixture random-effects regression model. In: Multivariate applications in substance use research: new methods for new questions. Hillsdale, NJ: Lawrence Erlbaum; 2000:79-112.

34. Sinharay S, Stern HS, Russell D. The use of multiple imputation for the analysis of missing data. Psychol Methods 2001 Dec;6(4):317-329. [Medline: 11778675]

35. Levav J, Fitzsimons GJ. When questions change behavior: the role of ease of representation. Psychol Sci 2006 Mar;17(3):207-213. [doi: 10.1111/j.1467-9280.2006.01687.x] [Medline: 16507060]

Abbreviations

CREMAS: Checklist for Reporting EMA Studies
EMA: ecological momentary assessment

MAR: missing at random
MCAR: missing completely at random
PA: physical activity
PDA: personal digital assistant
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Original Paper

"Less Than A Wife": A Study of Polycystic Ovary Syndrome Content in Teen and Women's Digital Magazines

Ninive Sanchez[1], MSW, MS, PhD; Hillary Jones[2], MSW, LICASW

[1]School of Social Work, University of Missouri, Columbia, MO, United States
[2]Catholic Community Services of Western Washington, Newcastle, WA, United States

Corresponding Author:
Ninive Sanchez, MSW, MS, PhD
School of Social Work
University of Missouri
712 Clark Hall
Columbia, MO, 65211
United States
Phone: 1 573 882 0920
Fax: 1 573 882 8926
Email: sanchezni@missouri.edu

Abstract

Background: Polycystic ovary syndrome (PCOS) is a major public health problem that affects women's physical and mental health. According to the US National Institutes of Health Office of Disease Prevention, there is a need to improve public awareness of the syndrome among health care providers and the public. Women's magazines are a type of "edutainment" that publish health content in addition to beauty, fashion, and entertainment content. These media have the potential to expose primarily female readers to content on PCOS and influence readers' beliefs and attitudes about women with PCOS.

Objective: The objective of this study was to explore how digital (online) teen and women's magazines portray women with PCOS.

Methods: We used data from the Alliance for Audited Media to identify popular digital teen and women's magazines with circulation rates ≥1,000,001. We also included magazines with circulation rates 100,001–1,000,000 directed toward racial and ethnic minority readers. A search of magazine websites over a 1-month period in 2015 yielded 21 magazines (eg, Glamour, Cosmopolitan en Español, Essence, and O, The Oprah Magazine) and 170 articles containing "PCOS" and "polycystic ovary syndrome." Textual analysis using a grounded theory approach was used to identify themes.

Results: Articles depicted PCOS symptoms as a hindrance to women's social roles as wives and mothers and largely placed personal responsibility on women to improve their health. To a lesser extent, women were depicted as using their personal experience with PCOS to advocate for women's health. Experiences of Latina and African American women and adolescents with PCOS were absent from women's magazine articles.

Conclusions: The findings can inform health education programs that teach women to be critical consumers of PCOS-related content in digital women's magazines. Future research on PCOS content in digital teen and women's magazines can help researchers, patients, and consumer groups engage with the media to increase public awareness of PCOS.

Introduction

In the United States, 35% of adults, or about 1 in 3 Americans, have used the Internet to find out what medical condition they or someone else might have [1]. Younger age groups and women increasingly use the Internet and social media to access health-related information [2,3]. Among the people searching for health-related information online are women typing "PCOS," "polycystic ovaries syndrome," and "PCOS symptoms" [4].

Overview of PCOS

Polycystic ovary syndrome (PCOS) is a complex, chronic condition the US National Institutes of Health (NIH) Office of

Disease Prevention has described as a major public health problem for women in the United States. PCOS is characterized by a range of symptoms including irregular or no menstrual periods, excess hair growth on the face and body (hirsutism), weight gain, acne, ovarian cysts, and alopecia [5]. PCOS can also increase women's risk of type 2 diabetes mellitus [6], cardiovascular disease [7], infertility [8], anxiety and depression [9-12], and poor health-related quality of life [13].

PCOS can begin in adolescence [14] and worsen across the life course if untreated or poorly managed [15,16]. The prevalence of PCOS among adult women (18-45 years of age) in the United States is about 7% [17]. However, research with a community sample of women (27-34 years of age) in the United Kingdom suggests that the prevalence of PCOS may be even higher, from 7.8% to 20.6%, depending on which of 3 diagnostic criteria for PCOS is used. Additionally, using 2 of these criteria, NIH and Rotterdam, about 70% of women with PCOS symptoms are undiagnosed [18]. See Goodman et al [19] and Rosenfield [20] for a review of the diagnosis and treatment of PCOS in adolescents and adults.

The 2012 NIH Evidence-based Methodology Workshop on Polycystic Ovary Syndrome recommended establishing "multidisciplinary programs to improve public and health care provider awareness and management for women who currently have the syndrome" [5]. The lack of awareness about PCOS is partly due to the condition's misleading name, which suggests that it is a problem of the ovaries; however, the presence of polycystic ovaries alone does not indicate that a woman has PCOS [5]. Additionally, although PCOS is a highly prevalent condition, it does not have the "celebrity" status of other well-known conditions [21].

PCOS Online

Women with PCOS describe the process of searching for answers to their symptoms as emotional, confusing, frustrating, devastating [22], and exhausting [23]. In addition to seeing several doctors before receiving a diagnosis, women report diagnosing themselves by searching for information online [24]. Women may search the Internet for information on PCOS they cannot find or understand in books or other print sources [25], and when they do not receive sufficient information and support from health care providers [23]. In addition, the Internet provides women with a convenient, private, and accessible way to access information on symptoms such as obesity and hirsutism (excessive hair) that may be too embarrassing to discuss in person [25]. Women communicate and share experiences with PCOS via chat groups, email lists [25], online support groups [23], and social networking sites like Facebook and Twitter [4].

Research on online information about PCOS has primarily focused on accuracy of this coverage. Information on PCOS appears on social networking sites, government health and professional associations' websites, and nonprofit organizations' websites [4,26], and websites vary regarding PCOS symptoms, long-term health effects, and management strategies [4,27]. Websites also tend to omit information on authors, editorial review process, publication date, strength of evidence reported (eg, use of randomized controlled trials), PCOS treatment based

on practice guidelines [27], privacy and confidentiality of personal data, and sources [4].

Media and Health

Media do not just provide consumers with health information, however; they also create social constructions of health, illness, and medical care that have implications for how consumers view and manage their health [28]. Teen and women's magazines have traditionally conflated health and beauty [29,30]. For instance, an analysis of body-related content published from 1993 to 2003 in *Seventeen*, a magazine for adolescent girls, found that content created constructions of desirable bodies as lean, toned, and free of hair and acne [31]. Equating health with beauty can be problematic for readers with PCOS. For instance, teen magazines publish techniques such as hair removal strategies and quick-fix diet and exercise routines readers can follow to achieve the desirable body [31]. Hair removal techniques that do not address the underlying causes of excessive hair growth can mask the severity of PCOS-related hirsutism [32]. Diet and exercise routines to achieve the desirable body in a short amount of time can overlook the importance of long-term lifestyle changes in diet and physical activity needed to manage PCOS. On the other hand, content in women's magazines that encourages consumers to collaborate with health care providers, be cautious about treatments and medications, and look to friends and family for social support may help consumers manage their health [33].

Edutainment

Magazines are a type of entertainment-education or "edutainment," which refers to the placement of educational content in entertainment messages [34]. Magazines publish health content in addition to beauty, fashion, and entertainment content. Teen magazines have published health content on topics such as sexually transmitted infections, gynecological visits, and birth control [35], and women's magazines have published on breast cancer [36], heart disease, osteoporosis, and depression [33]. Interestingly, there is evidence in the literature that magazine readers come across content on PCOS. As an adolescent pointed out,

> So I'm sitting inside the waiting room and I'm reading through and looking through magazines, and I came across an article in the Seventeen magazine: polycystic ovary disease. And a light bulb went off in my head. I'm like, hey, I gained weight for no reason. I have hair all over my face and it's popping up in places I don't want the hair to pop up. And so, I took it to my gynecologist and she was like, "That's what you have" [22].

Research also suggests that women with PCOS relate information they see in the media to their health, as the following quote illustrates: "If you look at how the media portrays what is feminine, lack of body hair, thinness, and a beautiful complexion, you don't have these with PCOS and wish you did" [22].

While several health topics have been featured across a range of edutainment [34], researchers have not examined content on PCOS in teen and women's magazines. These media have the

XSL•FO

RenderX

potential to expose primarily female readers to content on PCOS and influence readers' beliefs and attitudes about women with PCOS. In particular, online magazines, also referred to as digital magazines, are capable of reaching large audiences due to the lower cost of production and distribution, compared with print magazines [37].

This study aimed to understand content on PCOS in digital teen and women's magazines. The following research questions guided this study: (1) Do digital teen and women's magazines publish articles on PCOS? (2) Which magazines publish articles on PCOS? (3) How do digital teen and women's magazines portray adolescents and women with PCOS? (4) What is the discourse and ideology related to health, illness, and gender in these media?

This research can provide insight into health communication on PCOS in digital magazines to which readers are exposed. The findings can inform health education programs that teach women to be critical consumers of PCOS-related content in digital women's magazines.

Methods

Magazine Selection

We used data from the Alliance for Audited Media (AAM) to identify the highest-circulating teen and women's magazines with digital editions published in the United States. The AAM is a not-for-profit organization founded in 1914 (formerly known as the Audit Bureau of Circulation) that provides audited circulation figures, among other services and data, for newspapers, magazines, and digital media companies in the United States and Canada [38]. Advertising firms use these services and data to obtain details on various media types (eg, magazines, newspapers) to most effectively reach target markets or audiences. AAM is also available to academic institutions with electronic subscriptions to AAM [39]. Researchers, for instance, have used AAM to study the portrayal of gun violence and serious mental illness in news media coverage [40].

Consumer magazines with membership to AAM file publishers' statements with AAM every 6 months. Publishers' statements are claims made by consumer magazines that contain information on the publication's field served and average circulation of print and digital issues. AAM defines field served as the publisher's description of the markets or occupations whose interest the editorial content is directed toward. In other words, field served describes the publication's target audience or readership. Additionally, AAM defines digital edition circulation as the distribution of a magazine's content via electronic means. The digital edition maintains the same identity (eg, the same name and logo characteristics) as the host (or print) publication [38].

First, we identified magazines for this study by selecting AAM's circulation range of ≥1,000,001. This range was used as an indicator of popular consumer magazines. Second, we retrieved publishers' statements for magazines that met the circulation criteria. Magazines were selected for inclusion in this study if the publisher's statement described the field served as women or female teens. This yielded a total of 24 magazines. We also

selected magazines with circulation rates 100,001–1,000,000 in an effort to include magazines catering to racial and ethnic minority readers that may not have had the highest circulation rates. Publishers' statements were also retrieved for these magazines, yielding 3 additional magazines whose target audience or readership was women or female teens.

Article Selection

We independently searched each of the magazine websites from January 22 to February 25, 2015, to identify articles containing the keywords "PCOS" and "polycystic ovary syndrome." We selected a 1-month search period to determine whether any content on PCOS was published in teen and women's digital magazines. We compared the websites and number of articles containing the keywords. This yielded a total of 21 magazines and a total of 170 articles containing the keywords. The magazines were *Better Homes and Gardens*, *Cosmopolitan*, *Cosmopolitan en Español*, *Essence*, *Family Circle*, *Fitness*, *Glamour*, *Good Housekeeping*, *Health*, *MORE*, *O, The Oprah Magazine*, *Parenting*, *Parents*, *Prevention*, *Redbook*, *SELF*, *Shape*, *Teen Vogue*, *Vanidades*, *Woman's Day*, and *Women's Health*; 2 of these magazines, *Essence* and *Teen Vogue*, did not contain the keywords in articles, but rather in the reader comments.

Article Coding and Textual Analysis

We imported the final 170 articles into NVivo 10 for Windows (QSR International) using the NCapture for NVivo add-on in Chrome (Google). NCapture is a Web browser extension that captures Web content (eg, screenshots of webpages) that can then be imported into NVivo 10 for Windows as PDF sources [41]. In 1 instance, there was a video embedded in a Redbook article that had no text in the article except for the article's title and caption. The video was a total of 2:51 minutes in length and was streamed directly on Redbook's site, rather than taking the reader to a video-sharing website such as YouTube (YouTube, LLC). We transcribed the video and included the content in the analysis.

The 2 researchers independently coded each of the 170 articles to identify the following 5 primary content areas published in women's magazines: medical advice, letters or advice, personal stories, advertisements, and visual imagery. Medical advice pieces discussed PCOS in terms of warning signs, treatment, and prognosis. Letters or advice columns were those where questions about PCOS were asked and answered. Personal stories or narratives were based on women's own experiences with PCOS. Advertisements included health products and services related to PCOS. Visual imagery included images associated with the articles. We focused on the articles containing medical advice, letters or advice, and personal stories or narratives.

The 2 independent coders, both women, conducted a textual analysis using a grounded theory approach to coding to identify themes. The lead researcher identified as Mexican American and had background in psychology, social work, and public health. The graduate student research assistant, who was unfamiliar with the study aims, identified as white and had a

background in social work and public health. The lead researcher was fluent in Spanish and the research assistant was proficient.

During the data collection process, each researcher kept memos in which she documented the date the magazine article was retrieved and the titles of the magazine and article. Reflective information such as thoughts and questions that arose during this process were also noted.

Articles were read repeatedly and inductively, and new themes were noted as they emerged. The units of analysis were the article, title, caption, and user comments. Each researcher independently conducted an initial, unrestricted coding of the text in magazine articles by reading the texts line by line. Then, each researcher developed a coding scheme based on her interpretation of the text and memos. The researchers read the magazine articles as cultural artifacts: as texts that provide insight into social practices, representations, and assumptions about society [42].

We held peer debriefings to compare coding, understand each other's interpretations of the data, integrate categories, and refine the codebook. The codebook contained the thematic categories, examples of text for each category, and codes associated with each category. During these meetings, we viewed the screenshots of the magazine articles and reviewed the memos to contextualize their interpretations. This approach follows Boellstorff and colleagues' recommendation that data collected in virtual spaces (eg, screenshots and chat logs) be used in conjunction with field notes, not as a substitute for field notes [43].

Finally, we retrieved magazine media kits from publishers' websites to obtain magazine readers' demographic characteristics [44-63]. A media kit is a resource created by a publisher to help prospective ad buyers evaluate advertising opportunities [64].

This study used publicly available data and was deemed exempt from review by the University of Michigan Institutional Review Board.

Results

This study examined 3 main themes and 8 subthemes related to women's social roles and responsibilities, personal responsibility to improve health, and use of personal experience with PCOS to advocate for women's health. This paper contains limited portions of publicly available magazine articles associated with the themes.

Table 1 lists the digital magazines included in the final sample, descriptions of the field served, total number of articles containing the keywords, and readership age and race or ethnicity. *Health*, *Parents*, *Women's Health*, *Prevention*, *Shape*, and *Women's Day* contained the most articles on PCOS.

Table 1. Publishers' statements (January 30, 2014 to June 30, 2014 statement period) in magazines with articles on polycystic ovary syndrome (PCOS) and readership characteristics.[a]

Magazine	Field served	No. of articles on PCOS	Median age of readership (years)	Race/ethnicity of readership
Better Homes and Gardens	"*BETTER HOMES AND GARDENS* inspires women with infinite possibilities for creativity and self-expression. Each issue delivers smart, approachable editorial on design and individual style, decorating and gardening, food and entertaining, and personal and family well-being."	5	51.02	NA[b]
Cosmopolitan	"*COSMOPOLITAN* focuses on personal growth, relationships and careers, with expanded reporting on fashion and beauty, health and fitness. Covered as well are celebrities and pop culture...and just about everything else women want to know about."	3	34.7	NA
Cosmopolitan en Español [c]	"The assertive magazine for the independent Latin woman. Forward. Innovative. Successful. The Cosmo woman is all this and much more! *COSMOPOLITAN EN ESPAÑOL*, published as part of a joint venture with the Hearst Corporation, helps its readers to successfully balance their professional and personal lives. Editorial emphasis is on beauty, fashion and looking sensational."	2	42.1	NA
Essence	"A lifestyle magazine for today's African-American Woman."	1	42.1	NA
Family Circle	"*FAMILY CIRCLE* speaks to moms of tweens and teens. *FAMILY CIRCLE* delivers advice for tough parenting challenges; provides suggestions for family activities; offers quick and healthy family recipes; and showcases projects to create a comfortable home. *FAMILY CIRCLE* features the latest health, diet, fitness and style news; and offers beauty and fashion tips."	3	55	NA
Fitness	"A women's magazine addressing fitness as a lifestyle."	6	40	NA
Glamour	"Every issue of *GLAMOUR* includes news-making coverage of beauty, fashion, health and relationships as well as women's issues, work, money and more."	1	35.6	NA
Good Housekeeping	"Woman, Her Home and Her Family."	4	55.5	White: 12,940 (83.5%); Black/African American: 1588 (10.2%); Spanish/Hispanic Origin: 992 (6.4%)
Health	"*HEALTH* is the magazine for women who have discovered a new kind of healthy living. It provides information and inspiration on all aspects of healthy living—from cutting-edge health ideas to food, fitness, beauty and relationships."	30	49	NA
MORE	"*MORE* magazine is a fashion, beauty, trend and health guide for women of influence."	4	54	NA
O, The Oprah Magazine "A Desirable Audience"	"*O, THE OPRAH MAGAZINE* covers 360 degrees of a woman's life, from fashion and beauty, to relationships, food, home design, books, health and fitness, work and finance, technology, self-discovery and caring for others. The magazine encourages the reader to embrace her life, with the goal of becoming more of who she really is."	4	50.6	African Americans: 4095 (34%)

Magazine	Field served	No. of articles on PCOS	Median age of readership (years)	Race/ethnicity of readership
Parents	"*PARENTS*—the magazine mothers with young children turn to for guidance and information needed to raise happy, healthy, well-adjusted children."	24	34.9	NA
Parenting [c,d]	"Reality-tested ideas and support for moms, by moms."	11	NA	NA
Prevention	"Women's magazine with interest in good health and personal fitness. Editorial focus on healthful foods, meal planning and preparation, nutrition, skin and body care, self-improvement and other topics contributing to a healthful lifestyle."	16	57	NA
Redbook [e]	"*REDBOOK* makes great style accessible for women aged 25 to 54. Editorial coverage includes beauty, fashion, home décor, fitness and nutrition, money management, relationships and personal growth."	6	Women aged 35–49: 1722 (28.6%)	NA
SELF	"*SELF* is the magazine about the healthy well-being of the whole woman—body, mind and emotions. It encompasses everything that makes her unique, from her passions to her personal style."	2	42	NA
Shape	"Young, educated, affluent women leading active lifestyles who use fitness, fashion, and beauty to be their best."	15	39.6	NA
Teen Vogue	"A fashion and beauty magazine for teen girls. With a focus on fashion, coverage also includes entertainment and music, health and beauty, and inspiring profiles."	1	27	NA
Vanidades	"A women's beauty, fashion and lifestyle magazine in the U.S. Hispanic market. *VANIDADES* addresses the myriad interests of today's woman. Beauty and fashion are at the forefront, featuring the most recent developments in cosmetology and the latest offerings from the most renowned fashion designers."	1	39	Country/region of birth: Mexico: 40%; South America: 12%; United States: 11%; Central America: 8%; Cuba: 6%; Dominican Republic: 4%; Puerto Rico: 3% other: 16%
Woman's Day	"Woman's Service Field."	10	55.5	NA
Women's Health	"The *WOMEN'S HEALTH* brand was created for the woman who sees being healthy—physically and emotionally—as her edge. Our voice is how women speak to each other—with a tone and look that is uniquely *WOMEN'S HEALTH*."	21	33	NA

[a]Readership age and race/ethnicity taken from magazine media kits. Readership demographics for digital (Web) issue audiences provided when available.

[b]NA: Not available.

[c]Media kit for *Cosmopolitan en Español* and *Parenting* unavailable.

[d]Publisher's statement only available for the January 30, 2013 to June 30, 2013 statement period.

[e]Median age unavailable.

Women's Social Roles and Responsibilities

PCOS was depicted as a barrier to childbearing, starting a family, and effective breastfeeding. In turn, these issues hindered women's roles as wives and mothers. A total of 28 instances of this theme emerged from 9 magazines.

Childbearing and Role as a Wife

Childbearing was depicted as a wife's responsibility. Women's inability to have children lessened women's worth as wives. *Essence* published the following in a piece titled *Steve Harvey Morning Show's Daily Strawberry Letter*, a radio show in which listeners submit letters online describing personal issues and

XSL•FO
RenderX

problems, in hopes that hosts Steve Harvey and Shirley Strawberry will give them advice:

> *Today's topic: Less Than A Wife*
>
> *Hello Steve, Shirley, and Morning Crew. I am a 39 yr old woman with a wonderful husband. He is a God fearing man, that is an excellent provider, and the best husband that any woman could ask for. He is the most unselfish person that I have ever met. We desperately want to start a family but this past summer I was diagnosed with diabetes and ended up in ICU [intensive care unit] with a blood sugar of 980...... I also have PCOS (Polysystic Ovary Snydrome [sic]) which is already an issue and now this. I feel like less than a wife because I can not give my husband the one thing that he wants the most. [65]*

Use of "wonderful" suggests the writer holds her husband in high regard. Her description of him as "an excellent provider" and as "the best husband that any woman could ask for" also suggest she perceives him as meeting his role and responsibilities as a husband. The writer's tone becomes somber as she blames her poor health for preventing pregnancy and PCOS as an additional barrier to starting the family she and her husband "desperately want."

Childbearing and Starting a Family

Women's primary motivation to lose weight and manage PCOS symptoms was to increase their likelihood of pregnancy and form a family. The following excerpt published in *Fitness* featured Olivia Ward, season 11 winner of *The Biggest Loser*, a reality weight-loss show:

> *Olivia, who suffers from polycystic ovary syndrome, a condition that causes ovarian cysts and often disrupts the menstrual cycle, went on the show because doctors told her she needed to lose weight if she hoped to improve her chances of getting pregnant. She's now looking forward to starting a family with Ben. [66]*

The excerpt indicates that Olivia was unable to have children as an obese woman with PCOS. PCOS is depicted as a barrier to starting a family with her husband.

Breastfeeding

In 2 instances, PCOS was depicted as a barrier to effective breastfeeding. In the following excerpt, *Parents* magazine responds to the following reader's question: "Is breastfeeding when overweight difficult?":

> *Most plus-size women can breastfeed just as well as other women. However, if you're overweight because of a hormone or endocrine problem, like polycystic ovary syndrome (PCOS) or thyroid disease, you may have issues with your milk supply, since these conditions can also affect production of the hormones that trigger your body to make milk. But having PCOS, thyroid disease, or other hormone-related illnesses doesn't necessarily mean you won't be able to breastfeed. You may just have to supplement more often with formula to keep your baby fully nourished.*

> *If you're worried, it's a good idea to talk to a lactation consultant to help you and your baby get the hang of nursing if you run into problems. [67]*

The reader is concerned that being overweight could affect breastfeeding. The piece depicts PCOS as a potential barrier to nursing if it limits milk supply, a barrier that can affect the mother's caregiving and the child's nutrition.

Personal Responsibility for One's Health

Women were depicted as responsible for improving their own health through diet, exercise, hard work, and seeking good-quality health care. A total of 109 instances of this theme emerged from 17 magazines.

Lifestyle Changes

Magazine articles largely placed personal responsibility on women to improve their health with lifestyle changes. Women's stories illustrated ways in which women changed their diet and exercise through hard work and dedication. The following piece, titled *12 crazy-inspiring photos and details of weight loss success stories*, appeared in *Women's Health*:

> *The Lifestyle: Courtney dove into daily workouts. She jumped rope, punched bags, and learned how to fight. She fueled up with veggies and lean protein, and gave up ice cream after learning to satisfy her sweet tooth with fruits like pineapple.*
>
> *The Reward: Boxing also gave Courtney a major self-esteem boost. "I used to be afraid of talking in front of people," she says, "but now I'm confident." And although her PCOS is permanent, she feels healthy and has tons of energy. "I know I can face anything," she says. "Just like in the ring, you can always go one more round. You just have to dig for it." [68]*

Courtney's boxing is a metaphor for her "fight" against PCOS, a chronic condition Courtney is confident she can "face" or manage each "round." Courtney's hard work and dedication in adopting a healthy lifestyle was akin to training for a boxing tournament, in which she learned to fight an opponent. In this case, poor diet and inactivity were the opponents or barriers to weight loss. "You just have to dig for it" suggests that Courtney views one's ability to change as a result of hard work.

Good-Quality Health Care

Articles also largely placed personal responsibility on women to seek good-quality health care. The following piece, titled *Could you have PCOS?*, appeared in *SELF*:

> *The key is to see a health care provider with knowledge of the condition (ask whether the doctor frequently sees patients with PCOS). You can also request a referral to an endocrinologist. And if you do have PCOS, remember: The condition can be controlled, but it takes discipline. [69]*

The piece places responsibility on women to find a knowledgeable health care provider and request referrals to specialists. Additionally, the piece reminds women that PCOS

can be controlled with "discipline," suggesting that self-control is a primary factor in managing health.

Personal Experience as a Catalyst for Advocacy

Women were portrayed as using their personal experience with PCOS to promote women's health by sharing knowledge and expertise, promoting acceptance of one's body, and creating dialogue about miscarriage and infertility. A total of 8 instances of this theme emerged from 5 magazines.

Sharing Knowledge and Experience

Personal stories depicted women's knowledge of and experience with PCOS symptoms as valuable. Women expressed a desire to share their knowledge and experience with other women.

The following is an excerpt of a YouTube video embedded in a *Redbook* article titled *"I'm giving infertility a voice."* Carla opens up about her struggle with polycystic ovarian syndrome. Carla begins by telling viewers "I wish I had known that infertility isn't something to be ashamed of." She goes on to say she is 28 years old and has been diagnosed with PCOS and factor V Leiden, a condition that can increase one's risk of developing abnormal blood clots [70]. Then, Carla reflects on her experience with PCOS-related infertility:

> *Infertility can be really lonely. Besides the shame you might feel, there's a lot of misconception and misinformation floating around. I made it my mission to combat that and give infertility a voice by sharing my journey and educating those around me. I'm really proud of myself that I can look back on my journey to motherhood and know that not only did I become more resilient as an individual, and not only did it strengthen my marriage, but I was able to give hope to other people by educating them about infertility. I wish I hadn't been so ashamed about my diagnosis, but I'm glad I eventually realized, it's not my fault. [71]*

Carla characterizes infertility as a lonely and shameful experience. The video suggests that Carla experienced shame as a result of her inability to conceive and fulfill her gender roles as a woman, mother, and wife. The video also suggests that Carla's distressing experiences with infertility created a sense of duty and responsibility to "give hope" to women struggling with infertility. The excerpt concludes with Carla expressing happiness about realizing she was not to blame for her diagnosis of infertility, a rare contrast to the discourse on personal responsibility for one's health.

Promoting Acceptance

Encouraging women to accept their bodies was another way in which articles portrayed women as advocates for women's health. For instance, *Health* published a piece featuring Whitney Thore, the star of the reality TV show *My Big Fat Fabulous Life*, and her efforts to embrace her body [72]. Whitney gained notoriety when she posted a YouTube video titled *A Fat Girl Dancing*, in which she danced to a popular pop song [73]. According to *Health*:

> *Her videos became so popular that the now 30-year-old Thore launched the*

> *#NoBodyShameCampaign, a movement that encourages self-love and acceptance no matter your size or gender.*

> *Even Thore admits her weight gain wasn't due solely to PCOS. "Now I've gained 200 pounds so certainly I take personal responsibility. That's not all because of a medical condition," Thore told the Today Show last year. "But definitely I felt so ashamed once I started gaining weight that I was too embarrassed to even go to a doctor." If you think you've put on too much weight, it's important to take the time to get it checked out. [74]*

Thore used her knowledge of and experience with PCOS, as well as her popularity, to launch a campaign that promotes love and acceptance of one's body. Thore's campaign challenges social norms of what ideal bodies should look like. Use of "personal responsibility" suggests Thore believes personal choices, perhaps lack of self-discipline and control, contributed to her weight gain.

Creating Dialogue

Women's magazines also featured personal stories of celebrities who, unlike Thore, were thin and had no visible signs of PCOS. This included a piece in *Parents* magazine titled *Jaime King opens up: miscarriage, IVF—and then natural conception* featuring actress and model Jaime King:

> *The 35-year-old actress speaks openly about the traumatic process in this week's People magazine, revealing that she had five miscarriages, endured five rounds of IVF [in vitro fertilization], and 26 rounds of IUI [intrauterine insemination]. She says that a diagnosis of endometriosis and polycystic ovary syndrome were behind her difficulties.*

> *So with a healthy and happy 1-year-old baby in her life, and that suffering behind her, why is she talking about her miscarriage experience now? Because she says she wants to open up lines of dialogue about a topic that can be stubbornly closed for suffering hopeful moms-to-be out there. "I was hiding what I was going through for so long, and I hear about so many women going through what I went through," she told the magazine. "If I'm open about it, hopefully it won't be so taboo to talk about it." [75]*

King publicly discloses her PCOS diagnosis as an underlying cause of her infertility. "Traumatic" suggests that King's experiences with miscarriage and forms of assisted reproductive technology such as in vitro fertilization and intrauterine insemination negatively affected her physical and emotional well-being. King's motivation for publicly sharing her experience is to "open up lines of dialogue" about infertility and facilitate dialogue about a topic socially unacceptable or forbidden to discuss publicly.

Discussion

To our knowledge, this study is the first to explore PCOS content in digital teen and women's magazines. The presence of PCOS content in digital magazines in this study counters

previous findings that women's magazines tend to feature health content that is "old news," that is, information the public is generally familiar with [33]. PCOS is far from being old news to health care providers, patients, and the general public. Indeed, PCOS is a topic that within the last few years medical experts have referred to as a "hot topic" and "new frontier" in the area of female fertility [76].

This study found that digital magazines depicted PCOS as a barrier to having children that hindered women's roles as wives and mothers. These findings are consistent with Williams and colleagues' findings that women perceive PCOS as interrupting their plans to start a family and the timing of starting a family [23,77]. The portrayal of PCOS as a possible barrier to breastfeeding in women's magazines supports nascent research that suggests that women living with PCOS experience breastfeeding challenges. These challenges include low breast milk supply, limited ability to bond with their children through breastfeeding, and perceptions that breastfeeding can increase their daughters' risk for developing PCOS [78].

Portrayals of adolescents with PCOS were largely absent from magazine articles included in this study. This may be because teen magazines such as *Teen Vogue* primarily focus on clothing, shopping, and cosmetics [79]. In addition, the underrepresentation of adolescents in magazine articles may be associated with the limited research on PCOS in adolescence and its impact on health and mental health across the life course. Lack of research in these areas is one of the reasons why the NIH Evidence-based Methodology Workshop on Polycystic Ovary Syndrome has recommended conducting basic and translational research on PCOS in adolescence and the long-term consequences of PCOS diagnosed in adolescence [5]. This lack of research, combined with the challenges of diagnosing PCOS in adolescence [20], and the confusion surrounding the name "PCOS" [80,81] can affect the extent to which research trickles down to media such as magazines.

The absence of discourse on race and ethnicity, and the absence of PCOS content in magazines directed toward Latinas and African American women, suggests that portrayals of women with PCOS largely focus on the white body. This is cause for concern because Latinas and African American women are particularly at risk for PCOS due to high rates of obesity and metabolic problems (eg, insulin resistance, hypertension) [82] among these groups, and screening rates for metabolic problems among racial and ethnic minority women with PCOS tend to be low [83]. Interestingly, this study did not find articles depicting stories of African American women with PCOS in *O, The Oprah Magazine*, a magazine whose media kit claims to reach more African American readers than *Glamour*, *Redbook*, *SELF*, and *MORE* [54]. In addition, this study did not find articles depicting Latina or Spanish-speaking women's experiences with PCOS in either *Cosmopolitan en Español* or *Vanidades*, magazines published in Spanish and directed toward Latinas. The findings suggest that while discussion on race, ethnicity, and PCOS is not absent from the medical literature, it is absent in this type of popular media.

Magazine articles included in this study placed considerable demands on women to take responsibility for their own health.

The discourse on personal responsibility has been observed in other research with women's magazines. For instance, a review of readers' letters to women's health magazines found that writers tended to attribute their successes and failures in managing their health to individual behavior [30]. In addition, a content analysis of over 400 editorials with content on diet, overweight, and obesity published between 1984 and 2004 in mainstream and African American women's magazines found that these editorials primarily provided readers with strategies to change individual behavior such as reducing fast food intake and eating smaller portions [84]. Personal responsibility is certainly important in managing one's health. Indeed, women with PCOS report that diet and exercise are critical in controlling their weight and other PCOS symptoms [77]. However, it is important to keep in mind that personal responsibility discourse that blames individuals for poor lifestyle choices and poor health can perpetuate stigma and discrimination toward obese individuals and hinder efforts that address environmental and structural barriers to health [85,86].

The discourse on economic and environmental barriers to health was largely absent from teen and women's magazines. These barriers include high levels of food toxins and environmental exposure to industrial products (eg, bisphenol A in plastics) that appear to play a role in the development of PCOS [87], although further research in this area is needed. Additional barriers to health include lack of insurance coverage for infertility-based treatments and assisted reproductive technologies [88], neighborhood crime that limits one's ability to exercise in the community, and limited access to fresh fruits and vegetables [84].

Magazine articles depicted women as using their personal experiences with PCOS to be advocates for women's health. These women attached their names, voices, faces, bodies, emotions, and personalities to PCOS by making their most personal, private experiences with PCOS (eg, infertility, obesity, and poor body image) public. These women appear to have engaged in meaning making, a process in which individuals try to make sense of or comprehend a stressful event and search for the value or significance of the event in their lives. Meaning making can result in acceptance of the event, personal growth, enhanced coping skills, a greater appreciation for life, and a change in identity [89]. In their journey with PCOS, women gained power and control over their body and their lives, and realized that the knowledge and experience they gained during their journey was valuable and should be shared with others to educate them about complex health issues, give others hope, and facilitate dialogue about taboo topics.

Study Limitations

A limitation of this study was the 1-month search period that did not lend itself to the study of potential changes in PCOS content over time. Articles on PCOS were included in the study if they were available online during the 1-month search period. As a result, the articles included in this study may not be representative of PCOS content published in these media over a longer period of time. Despite this limitation, the 1-month search period was effective in demonstrating that, indeed, women's digital magazines publish content on PCOS.

This study also found that, in some cases, articles were published online months and even years preceding the search period. This illustrates how commercial publications such as magazines that once had short shelf lives as hard copies can continue to be accessible online for lengthy periods of time [90].

A second limitation was the method used to search for articles using the keywords "PCOS" and "polycystic ovary syndrome." We used only these keywords in this study. This may have excluded articles containing other names for PCOS such as "polycystic ovarian syndrome." In addition, the search method may not have yielded all articles containing the keywords. For instance, an article titled *7 celebrities who manage life with chronic pain* published in November 2011 in *Prevention* magazine containing the keywords was not returned by the magazine website's search engine using this search method. As a result, we did not include this article in this analysis. This indicates the search bars in digital magazines may not have been sensitive enough to find all articles containing the keywords. Future research should include a more rigorous search method.

Future research should also examine how readers respond to PCOS content in digital teen and women's magazines. This includes ways in which readers respond to women's personal stories and manage their health, or not, as a result of exposure to this content. Research in this area is especially important because the number of Internet users of online content to make health and health care decisions is rising [91].

Conclusion

This study highlights how digital teen and women's magazines can be used to understand health communication about PCOS. Research on popular media content is not readily available in the literature on PCOS, as much of the information on women with the condition comes from medical sources. This study found a lack of representation of adolescents and racial and ethnic minority women with PCOS in magazine articles. In addition, this study found social values and beliefs about women with PCOS embedded in the articles. The findings can inform health education programs that teach women to be critical consumers of PCOS-related content in digital women's magazines. Critical consumers have the skills to recognize how content is socially constructed and make informed decisions about their health through reflection, questioning, and analysis of media [92]. Future research on PCOS content in digital teen and women's magazines can help researchers, patients, and consumer groups engage with the media to increase public awareness of PCOS.

Conflicts of Interest

None declared.

References

1. Fox S, Duggan M. Health online 2013. Washington, DC: Pew Internet & American Life Project; 2013 Jan 15. URL: http://www.pewinternet.org/files/old-media//Files/Reports/PIP_HealthOnline.pdf [accessed 2016-03-30] [WebCite Cache ID 6gOiayEJG]

2. Atkinson NL, Saperstein SL, Pleis J. Using the internet for health-related activities: findings from a national probability sample. J Med Internet Res 2009;11(1):e4 [FREE Full text] [doi: 10.2196/jmir.1035] [Medline: 19275980]

3. Chou WS, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. J Med Internet Res 2009;11(4):e48 [FREE Full text] [doi: 10.2196/jmir.1249] [Medline: 19945947]

4. Mousiolis A, Michala L, Antsaklis A. Polycystic ovary syndrome: double click and right check. What do patients learn from the Internet about PCOS? Eur J Obstet Gynecol Reprod Biol 2012 Jul;163(1):43-46. [doi: 10.1016/j.ejogrb.2012.03.028] [Medline: 22512829]

5. National Institutes of Health. Evidence-based Methodology Workshop on Polycystic Ovary Syndrome. 2012. URL: https://prevention.nih.gov/docs/programs/pcos/FinalReport.pdf [accessed 2015-11-30] [WebCite Cache ID 6dR0OLGHh]

6. Gambineri A, Patton L, Altieri P, Pagotto U, Pizzi C, Manzoli L, et al. Polycystic ovary syndrome is a risk factor for type 2 diabetes: results from a long-term prospective study. Diabetes 2012 Sep;61(9):2369-2374 [FREE Full text] [doi: 10.2337/db11-1360] [Medline: 22698921]

7. Wild S, Pierpoint T, Jacobs H, McKeigue P. Long-term consequences of polycystic ovary syndrome: results of a 31 year follow-up study. Hum Fertil (Camb) 2000;3(2):101-105. [Medline: 11844363]

8. Amsterdam ESHRE/ASRM-Sponsored 3rd PCOS Consensus Workshop Group. Consensus on women's health aspects of polycystic ovary syndrome (PCOS). Hum Reprod 2012 Jan;27(1):14-24 [FREE Full text] [doi: 10.1093/humrep/der396] [Medline: 22147920]

9. Annagür BB, Tazegül A, Uguz F, Kerimoglu OS, Tekinarslan E, Celik C. Biological correlates of major depression and generalized anxiety disorder in women with polycystic ovary syndrome. J Psychosom Res 2013 Mar;74(3):244-247. [doi: 10.1016/j.jpsychores.2013.01.002] [Medline: 23438716]

10. Dokras A. Mood and anxiety disorders in women with PCOS. Steroids 2012 Mar 10;77(4):338-341. [doi: 10.1016/j.steroids.2011.12.008] [Medline: 22178257]

11. Dokras A, Clifton S, Futterweit W, Wild R. Increased prevalence of anxiety symptoms in women with polycystic ovary syndrome: systematic review and meta-analysis. Fertil Steril 2012 Jan;97(1):225-30.e2. [doi: 10.1016/j.fertnstert.2011.10.022] [Medline: 22127370]

12. Livadas S, Chaskou S, Kandaraki AA, Skourletos G, Economou F, Christou M, et al. Anxiety is associated with hormonal and metabolic profile in women with polycystic ovarian syndrome. Clin Endocrinol (Oxf) 2011 Nov;75(5):698-703. [doi: 10.1111/j.1365-2265.2011.04122.x] [Medline: 21605157]

13. Li Y, Li Y, Yu Ng EH, Stener-Victorin E, Hou L, Wu T, et al. Polycystic ovary syndrome is associated with negatively variable impacts on domains of health-related quality of life: evidence from a meta-analysis. Fertil Steril 2011 Aug;96(2):452-458. [doi: 10.1016/j.fertnstert.2011.05.072] [Medline: 21703610]

14. Connor EL. Adolescent polycystic ovary syndrome. Adolesc Med State Art Rev 2012 Apr;23(1):164-77, xii. [Medline: 22764561]

15. Goverde AJ, Westerveld HE, Verhulst SM, Fauser BC. Polycystic ovary syndrome as a developmental disorder. Expert Review of Obstetrics & Gynecology 2008 Nov;3(6):775-787. [doi: 10.1586/17474108.3.6.775]

16. Sanchez N. A life course perspective on polycystic ovary syndrome. Int J Womens Health 2014 Jan;6:115-122 [FREE Full text] [doi: 10.2147/IJWH.S55748] [Medline: 24489477]

17. Azziz R, Woods KS, Reyna R, Key TJ, Knochenhauer ES, Yildiz BO. The prevalence and features of the polycystic ovary syndrome in an unselected population. J Clin Endocrinol Metab 2004 Jun;89(6):2745-2749. [doi: 10.1210/jc.2003-032046] [Medline: 15181052]

18. March WA, Moore VM, Willson KJ, Phillips DI, Norman RJ, Davies MJ. The prevalence of polycystic ovary syndrome in a community sample assessed under contrasting diagnostic criteria. Hum Reprod 2010 Feb;25(2):544-551 [FREE Full text] [doi: 10.1093/humrep/dep399] [Medline: 19910321]

19. Goodman NF, Cobin RH, Futterweit W, Glueck JS, Legro RS, Carmina E. American Association of Clinical Endocrinologists, American College of Endocrinology, and Androgen Excess and PCOS society disease state clinical review: guide to the best practices in the evaluation and treatment of polycystic ovary syndrome - part 1. Endocr Pract 2015 Nov;21(11):1291-1300. [doi: 10.4158/EP15748.DSC] [Medline: 26509855]

20. Rosenfield RL. The diagnosis of polycystic ovary syndrome in adolescents. Pediatrics 2015 Dec;136(6):1154-1165. [doi: 10.1542/peds.2015-1430] [Medline: 26598450]

21. Corley EA. A use-and-transformation model for evaluating public R&D: illustrations from polycystic ovarian syndrome (PCOS) research. Eval Program Plann 2007 Feb;30(1):21-35. [doi: 10.1016/j.evalprogplan.2006.09.001] [Medline: 17689311]

22. Snyder BS. The lived experience of women diagnosed with polycystic ovary syndrome. J Obstet Gynecol Neonatal Nurs 2006;35(3):385-392. [doi: 10.1111/j.1552-6909.2006.00047.x] [Medline: 16700688]

23. Williams S, Sheffield D, Knibb RC. 'Everything's from the inside out with PCOS': exploring women's experiences of living with polycystic ovary syndrome and co-morbidities through SkypeTM interviews. Health Psychology Open 2015;2(2):1-10. [doi: 10.1177/2055102915603051]

24. Crete J, Adamshick P. Managing polycystic ovary syndrome: what our patients are telling us. J Holist Nurs 2011 Dec;29(4):256-266. [doi: 10.1177/0898010111398660] [Medline: 21357181]

25. Avery JC, Braunack-Mayer AJ. The information needs of women diagnosed with Polycystic Ovarian Syndrome--implications for treatment and health outcomes. BMC Womens Health 2007;7:9 [FREE Full text] [doi: 10.1186/1472-6874-7-9] [Medline: 17578583]

26. Dean EH. An analysis of online resources for women with polycystic ovary syndrome. J Consumer Health Internet 2011 Oct;15(4):361-369. [doi: 10.1080/15398285.2011.623580]

27. Mallappa Saroja CS, Hanji Chandrashekar S. Polycystic ovaries: review of medical information on the internet for patients. Arch Gynecol Obstet 2010 May;281(5):839-843. [doi: 10.1007/s00404-010-1378-4] [Medline: 20143080]

28. Kline KN. Popular media and health. In: Thompson TL, Parrott R, Nussbaum JF, editors. The Routledge Handbook of Health Communication, 2nd edition. New York, NY: Routledge; 2011:252-267.

29. Madden H, Chamberlain K. Nutritional health messages in women's magazines: a conflicted space for women readers. J Health Psychol 2004 Jul;9(4):583-597. [doi: 10.1177/1359105304044044] [Medline: 15231058]

30. Newman C. Reader letters to women's health magazines. Feminist Media Stud 2007 Jun;7(2):155-170. [doi: 10.1080/14680770701287027]

31. Ballentine LW, Ogle JP. The making and unmaking of body problems in Seventeen magazine, 1992-2003. Fam Consumer Sci Res J 2005 Jun 01;33(4):281-307. [doi: 10.1177/1077727X04274114]

32. Keegan A, Liao L, Boyle M. 'Hirsutism': a psychological analysis. J Health Psychol 2003 May;8(3):327-345. [doi: 10.1177/13591053030083004] [Medline: 14670212]

33. Barnett B. Health as women's work: a pilot study on how women's magazines frame medical news and femininity. Women Lang 2006;29(2):1-12.

34. Singhal A, Rogers EM. A theoretical agenda for entertainment--education. Commun Theory 2002 May;12(2):117-135. [doi: 10.1111/j.1468-2885.2002.tb00262.x]

35. Keller J. Feminist editors and the new girl glossies: rashionable feminism or just another sexist rag? Womens Stud Int Forum 2011 Jan;34(1):1-12. [doi: 10.1016/j.wsif.2010.07.004]

36. Andsager JL, Powers A. Framing women's health with a sense-making approach: magazine coverage of breast cancer and implants. Health Commun 2001;13(2):163-185. [doi: 10.1207/S15327027HC1302_3] [Medline: 11451103]

37. Santos Silva D. The future of digital magazine publishing. Inf Serv Use 2011;31:301-310. [doi: 10.3233/ISU-2012-0661]

38. Alliance for Audited Media. How to Use an AAM Consumer Magazine Publisher's Statement. 2013. URL: http://auditedmedia.com/media/166295/magazinehowto.pdf [accessed 2015-11-30] [WebCite Cache ID 6dR5ABcoe]

39. Pittsley K. Where to advertise media planning resources on the web. Coll Res Libr News 2014;75:76-100.

40. McGinty EE, Webster DW, Jarlenski M, Barry CL. News media framing of serious mental illness and gun violence in the United States, 1997-2012. Am J Public Health 2014 Mar;104(3):406-413. [doi: 10.2105/AJPH.2013.301557] [Medline: 24432874]

41. QSR International. NVivo 10 for Windows Add-ons. URL: http://www.qsrinternational.com/product [accessed 2015-11-30] [WebCite Cache ID 6dR5K5pQJ]

42. Brennen BS. Qualitative Research Methods for Media Studies. New York, NY: Routledge; 2013.

43. Boellstorff T, Nardi B, Pearce C, Taylor TL. Ethnography and Virtual Worlds: A Handbook of Method. Princeton, NJ: Princeton University Press; 2012.

44. Better Homes and Gardens. New York, NY: BHG; 2015 Dec 01. Adult Readers URL: http://bhgmarketing.com/wp-content/uploads/2015/01/Adults_MRI-Fall2014.pdf [accessed 2015-11-30] [WebCite Cache ID 6dR5UP6aI]

45. Cosmopolitan Demographic Profile. New York, NY: Hearst Corporation; 2014. URL: http://www.cosmomediakit.com/r5/showkiosk.asp?listing_id=4785154&category_code=demo&category_id=77109 [accessed 2015-03-29]

46. Essence 2015 Media Kit. New York, NY: Essence Communications Inc; 2015. URL: http://www.essence.com/sites/default/files/promotions/mk/2015%20ESSENCE%20MEDIA%20KIT_FINAL_v2.pdf [accessed 2015-03-29] [WebCite Cache ID 6dR61QW4W]

47. The Family Circle Reader. New York, NY: Meredith Corp; 2013. URL: http://www.meredith.com/mediakit/familycircle/production/2015/docs/FCReaderProfile_Women_2014.pdf [accessed 2015-11-30] [WebCite Cache ID 6dR6KhUtk]

48. The Fitness Audience. New York, NY: Meredith Corp; 2013. URL: http://www.meredith.com/mediakit/2014FITNESS_MediaKit.pdf [accessed 2015-11-30] [WebCite Cache ID 6dR6bI48h]

49. Glamour Web Media Kit. New York, NY: Conde Nast; 2015. URL: http://www.condenast.com/brands/glamour/media-kit/web/pdf [accessed 2015-11-30] [WebCite Cache ID 6dR76kRAB]

50. Good Housekeeping Media Kit. New York, NY: Hearst Communications Inc URL: http://www.ghmediakit.com/r5/home.asp [accessed 2016-04-11] [WebCite Cache ID 6ggwTTIqD]

51. Health. Health. New York, NY: Time Inc; 2015. URL: http://img2.timeinc.net/health/static/advertise-digital/pdfs/FullMediaKit.pdf [accessed 2016-04-11] [WebCite Cache ID 6gcdyv6TO]

52. MORE Reader Profile: 2014 IPSOS Affluent Survey. New York, NY: Meredith Corp; 2015. URL: http://www.meredith.com/mediakit/more/2015/pdfs/15_MORE_Stats_IPSOS.pdf [accessed 2015-11-30] [WebCite Cache ID 6dR8c3R39]

53. MORE Reader Profile: MRI Doublebase 2014 Women. New York, NY: Meredith Corp; 2015. URL: http://www.meredith.com/mediakit/more/2015/pdfs/15_MORE_Stats_MRI_ReaderProfile.pdf [accessed 2015-11-30] [WebCite Cache ID 6dR8ptWK7]

54. O, The Oprah Magazine: Media Kit. New York, NY: Hearst Corporation; 2014. URL: http://www.omediakit.com/r5/cob_page.asp?category_code=audi [accessed 2016-04-08] [WebCite Cache ID 6gcb8WebQ]

55. Parents: Adult Demographic Profile. New York, NY: Meredith Corp; 2015. URL: http://mereditdirectmedia.com/magazines/parents/ [accessed 2015-11-30] [WebCite Cache ID 6dR9NxQFw]

56. Prevention 2015 Media Kit: 2015 Full Circulation Demographic Profile. Emmaus, PA: Rodale, Inc; 2015. URL: http://www.prevention.com/sites/prevention.com/files/PVN_MediaKit_2015_0.pdf [accessed 2016-04-08] [WebCite Cache ID 6gcak7FyO]

57. Redbook: 2015 Key Audience.: Hearst Corporation; 2015. URL: http://www.redbookmediakit.com/hotdata/publishers/redbookme2011484/advertiser/8329533/3194598/Key%20Audience.pdf [accessed 2015-03-29] [WebCite Cache ID 6dRAGl6S5]

58. SELF Web Media Kit. New York, NY: Conde Nast; 2013. URL: http://www.condenast.com/brands/self/media-kit/web/pdf [accessed 2016-04-08] [WebCite Cache ID 6gcZgYwoS]

59. Shape Media Kit 2015. New York, NY: Meredith Corporation; 2014. 2015 URL: http://www.shape.com/sites/shape.com/files/media/shape_2015_media_kit_2.pdf [accessed 2015-11-30] [WebCite Cache ID 6dRAqnVBW]

60. Teen Vogue Demographics. New York, NY: Conde Nast; 2016. URL: http://www.condenast.com/brands/teen-vogue/media-kit/web [accessed 2016-04-08] [WebCite Cache ID 6gcRBZAvO]

61. Vanidades 2014 U.S. Hispanic Media Kit. Miami, FL: Televisa Publishing + Digital; 2014. URL: http://www.avanimedia.com/Media-Kits/2014%20Vanidades.pdf [accessed 2015-11-30] [WebCite Cache ID 6dRB89kKZ]

62. Woman's Day 2015 Reader Profile. New York, NY: Hearst Corporation; 2015. URL: http://www.womansdaymediakit.com/hotdata/publishers/womansdaymk/advertiser/8329732/3196694/WD_ReaderProfile_Fall2014.pdf [accessed 2015-03-29] [WebCite Cache ID 6dVij93z9]

63. Women's Health Media Kit 2015 Digital. Emmaus, PA: Rodale, Inc; 2014. 2015 Digital URL: http://www.womenshealthmag.com/sites/womenshealthmag.com/files/mediakit/WH-MediaKit-Online.pdf [accessed 2015-12-03] [WebCite Cache ID 6dViT9vtu]

64. Doyle C. Media kit. In: A Dictionary of Marketing, 3rd edition. Cary, NC: Oxford University Press; 2011.

65. Essence. New York, NY: Essence Communications Inc; 2010 Jan 18. Steve Harvey Morning Show's Daily Strawberry Letter URL: http://www.essence.com/2010/01/19/steve-harvey-morning-shows-daily-strawbe [accessed 2015-12-03] [WebCite Cache ID 6dVhOBaCT]

66. Haney L. Biggest Loser's Olivia Ward's Tips for a Fit Marriage. New York, NY: Meredith Corporation; 2015. URL: http://www.fitnessmagazine.com/weight-loss/tips/olivia-ward-marriage-weight-loss-tips/ [accessed 2015-12-03] [WebCite Cache ID 6dVh0606U]

67. Is Breastfeeding When Overweight Difficult?. New York, NY: Meredith Corporation; 2015. URL: http://www.parents.com/advice/babies/breastfeeding/is-breastfeeding-when-overweight-difficult/ [accessed 2015-12-03] [WebCite Cache ID 6dVfyTQf5]

68. Stearns C. 12 Crazy-Inspiring Photos and Details of Weight Loss Success Stories. Emmaus, PA: Rodale, Inc; 2012 Oct 22. URL: http://www.womenshealthmag.com/weight-loss/weight-loss-success-stories [accessed 2015-12-03] [WebCite Cache ID 6dVfWabP0]

69. Could you have PCOS?. New York, NY: Conde Nast; 2007 Jan 23. URL: http://www.self.com/wellness/health/2007/01/signs-of-polycystic-ovary-syndrome [accessed 2015-12-03] [WebCite Cache ID 6dVdbFrPq]

70. May Clinic Staff. Diseases and Conditions: Factor V Leiden. Rochester, MN: Mayo Foundation for Medical Education and Research; 2015 Jul 14. URL: http://www.mayoclinic.org/diseases-conditions/factor-v-leiden/basics/definition/con-20032637 [accessed 2015-12-03] [WebCite Cache ID 6dVYmFcJs]

71. Redbook. "I'm Giving Infertility a Voice". New York, NY: Hearst Communications, Inc; 2010 Oct 10. URL: http://www.redbookmag.com/body/pregnancy-fertility/videos/a12813/im-giving-infertility-a-voice-video-1211447043001/ [accessed 2015-12-03] [WebCite Cache ID 6dVUXq8Qk]

72. My Big Fat Fabulous Life: About Whitney. Silver Spring, MD: TLC URL: http://www.tlc.com/tv-shows/my-big-fat-fabulous-life/about-my-big-fat-fabulous-life/ [accessed 2015-12-03] [WebCite Cache ID 6dVUJKIeS]

73. Thore WW. Original - A Fat Girl Dancing: Wiggle (Jason Derulo ft Snoop Dogg). San Bruno, CA: YouTube; 2014 May 27. URL: https://www.youtube.com/watch?v=Hj-VmZYcUhQ [accessed 2015-12-03] [WebCite Cache ID 6dVU8TU0z]

74. Swalin R. The Self-Acceptance Lesson You Can Learn From Whitney Thore of 'My Big Fat Fabulous Life'. New York, NY: Health Media Ventures, Inc; 2015 Jan 13. URL: http://news.health.com/2015/01/13/the-self-acceptance-lesson-you-can-learn-from-whitney-thore-of-my-big-fat-fabulous-life/ [accessed 2015-12-03] [WebCite Cache ID 6dVTup3Ck]

75. Dubin A. Jaime King Opens Up: Miscarriage, IVF—and Then Natural Conception. New York, NY: Meredith Corporation URL: http://www.parents.com/blogs/everything-pregnancy/2015/01/14/everything-pregnancy/jaime-king-opens-up-miscarriage-ivf-and-then-natural-conception/ [accessed 2015-12-03] [WebCite Cache ID 6dVTYYVbw]

76. Cecconi S. Editorial [hot topic: new frontiers in female reproduction and fertility (executive guest editor: sandra cecconi)]. Curr Pharm Des 2012;18(3):231-232. [Medline: 22229561]

77. Williams S, Sheffield D, Knibb RC. A snapshot of the lives of women with polycystic ovary syndrome: A photovoice investigation. J Health Psychol 2014 Sep 9:1-13. [doi: 10.1177/1359105314547941] [Medline: 25205774]

78. Baker E. Experiences of women living with polycystic ovary syndrome (PCOS): breastfeeding challenges and need for breastfeeding support. 2015 Nov 01 Presented at: American Public Health Association Annual Meeting; October 31 - November 04, 2015; Chicago, IL.

79. Massoni K. Fashioning Teenagers: A Cultural History of Seventeen Magazine. Walnut Creek, CA: Left Coast Press; 2010.

80. Teede H, Gibson-Helm M, Norman RJ, Boyle J. Polycystic ovary syndrome: perceptions and attitudes of women and primary health care physicians on features of PCOS and renaming the syndrome. J Clin Endocrinol Metab 2014 Jan;99(1):E107-E111. [doi: 10.1210/jc.2013-2978] [Medline: 24178791]

81. Azziz R. Polycystic ovary syndrome: what's in a name? J Clin Endocrinol Metab 2014 Apr;99(4):1142-1145 [FREE Full text] [doi: 10.1210/jc.2013-3996] [Medline: 24491161]

82. Fauser BC, Tarlatzis BC, Rebar RW, Legro RS, Balen AH, Lobo R, et al. Consensus on women's health aspects of polycystic ovary syndrome (PCOS): the Amsterdam ESHRE/ASRM-Sponsored 3rd PCOS Consensus Workshop Group. Fertil Steril 2012 Jan;97(1):28-38.e25. [doi: 10.1016/j.fertnstert.2011.09.024] [Medline: 22153789]

83. Mott MM, Kitos NR, Coviello AD. Practice Patterns in screening for metabolic disease in women with PCOS of diverse race-ethnic backgrounds. Endocr Pract 2014 Sep;20(9):855-863. [doi: 10.4158/EP13414.OR] [Medline: 24641921]

84. Campo S, Mastin T. Placing the burden on the individual: overweight and obesity in African American and mainstream women's magazines. Health Commun 2007;22(3):229-240. [doi: 10.1080/10410230701626885] [Medline: 17967145]

85. Brownell KD, Kersh R, Ludwig DS, Post RC, Puhl RM, Schwartz MB, et al. Personal responsibility and obesity: a constructive approach to a controversial issue. Health Aff (Millwood) 2010;29(3):379-387 [FREE Full text] [doi: 10.1377/hlthaff.2009.0739] [Medline: 20194976]

86. Puhl RM, Heuer CA. Obesity stigma: important considerations for public health. Am J Public Health 2010 Jun;100(6):1019-1028. [doi: 10.2105/AJPH.2009.159491] [Medline: 20075322]

87. Diamanti-Kandarakis E. Novel insights into the pathophysiology of PCOS: the role of environmental toxins. In: Diamanti-Kandarakis E, Nader S, Panidis D, editors. Novel Insights into the Pathophysiology and Treatment of PCOS. 1st edition. London: Future Medicine Ltd; 2013:38-48.

XSL•FO

RenderX

88. Brezina PR, Shah AA, Myers ER, Huang A, DeCherney AH. How Obamacare will impact reproductive health. Semin Reprod Med 2013 May;31(3):189-197. [doi: 10.1055/s-0033-1336599] [Medline: 23609149]

89. Park CL. Making sense of the meaning literature: an integrative review of meaning making and its effects on adjustment to stressful life events. Psychol Bull 2010 Mar;136(2):257-301. [doi: 10.1037/a0018301] [Medline: 20192563]

90. Lindlof TR, Taylor BC. Qualitative Communication Research Methods. Los Angeles, CA: SAGE; 2011.

91. Rainie L. The rise of the e-patient: understanding social networks and online health information seeking. Washington, DC: Pew Internet & American Life Project; 2012 Jan 12. URL: http://www.pewinternet.org/files/old-media/ /Files/Presentations/2012/Jan/2012%20-%201%2012%2012%20-%20Rise%20of%20e-patients%20-%20Providence%20St%20%20Joseph %20Medical%20CenterPDF.pdf [accessed 2015-12-03] [WebCite Cache ID 6dVR3JOFg]

92. Higgins JW, Begoray D. Exploring the borderlands between media and health: conceptualizing 'critical media health literacy'. J Media Lit Educ 2012;4(2):136-148.

Abbreviations

AAM: Alliance for Audited Media
NIH: National Institutes of Health
PCOS: polycystic ovary syndrome

XSL•FO

RenderX

Original Paper

Primary Care Provider Views About Usefulness and Dissemination of a Web-Based Depression Treatment Information Decision Aid

Julie Beaulac[1,2], PhD; Robin Westmacott[3], PhD; John R Walker[3], PhD; Gohar Vardanyan[4], MA; The Mobilizing Minds Research Group[4]

[1]The Ottawa Hospital, Psychology Department, Ottawa, ON, Canada

[2]The Ottawa Hospital Research Institute, Ottawa, ON, Canada

[3]The University of Manitoba, Department of Clinical Health Psychology, Winnipeg, MB, Canada

[4]The University of Manitoba, Winnipeg, MB, Canada

Corresponding Author:
Julie Beaulac, PhD
The Ottawa Hospital
Psychology Department
501 Smyth Road, Room 7300 General Campus
Ottawa, ON, K1H 8L6
Canada
Phone: 1 613 737 8899 ext 75078
Fax: 1 613 737 8895
Email: jbeaulac@toh.on.ca

Abstract

Background: Decisions related to mental health are often complex, problems often remain undetected and untreated, information unavailable or not used, and treatment decisions frequently not informed by best practice or patient preferences.

Objective: The objective of this paper was to obtain the opinions of health professionals working in primary health care settings about a Web-based information decision aid (IDA) for patients concerning treatment options for depression and the dissemination of the resources in primary care settings.

Methods: Participants were recruited from primary care clinics in Winnipeg and Ottawa, Canada, and included 48 family physicians, nurses, and primary care staff. The study design was a qualitative framework analytic approach of 5 focus groups. Focus groups were conducted during regular staff meetings, were digitally recorded, and transcripts created. Analysis involved a content and theme analysis.

Results: Seven key themes emerged including the key role of the primary care provider, common questions about treatments, treatment barriers, sources of patient information, concern about quality and quantity of available information, positive opinions about the IDA, and disseminating the IDA. The most common questions mentioned were about medication and side effects and alternatives to medication. Patients have limited access to alternative treatment options owing to cost and availability.

Conclusions: Practitioners evaluated the IDA positively. The resources were described as useful, supportive of providers' messages, and accessible for patients. There was unanimous consensus that information needs to be available electronically through the Internet.

Introduction

Informed decision-making is essential to good patient health outcomes. Patients are more likely to initiate and continue treatment when informed about choices [1], facilitating a more cost-effective use of health care resources [2]. Shared decision making is associated with improved quality of the decision in terms of knowledge and values and improvements in treatment progress [3]. For depression, higher participation is associated with improved treatment adherence and health outcomes [4-6].

One tool for facilitating informed decision making on health is the information decision aid (IDA), which engages patients in the decision-making process and helps them to make choices among different treatment options. Information decision aids present information about a condition, associated medical tests and treatment options, and the probabilities of risks and benefits of the different options to assist patients with the decision-making process [7]. Information decision aids can be used as self- or practitioner-administered tools, alone or as part of structured counseling or patient education [8]. The importance of IDAs as a tool to support decision making brings forward the question of their effective dissemination in primary health care settings.

Attitudes of clinicians toward Web-based decision aids and Web-based tools to guide the treatment process have been found to be favorable, including for depression [9] and other mental health conditions (eg, psychosis; [10]). However, challenges have been identified related to implementing tools in actual practice [9,10]. A recent review indicated that adoption of IDAs in clinical practice has been very limited [11], possibly because of clinician concern about the content of the decision aids and about how well the use of IDAs fits into the traditional workflow in primary care settings. The International Patient Decision Aid Standards (IPDAS) criteria [12] emphasize the importance of the involvement of clinicians in real-world settings in evaluating and planning for the implementation of IDAs [13], and more engagement of clinicians may improve implementation [14].

Primary care professionals have important potential in disseminating IDAs about many issues including mental health. There is a gap in knowledge related to dissemination of Web-based tools for mental health, particularly from the perspective of health providers. To effectively disseminate such tools, we need to understand where health professionals currently go for information, how they provide information to patients, and how resources can be made available to better assist them in providing care for common mental health problems such as depression.

There are few IDAs available in the mental health area [5,6]. The ones that are available tend to focus on medication treatment rather than the wider range of treatment options that are available to consumers [15]. Our team, consisting of mental health professionals and young adults interested in mental health, has developed an IDA for the Canadian public focused on treatment of depression. The development of the IDA was guided by the IPDAS criteria [12]. The developers did extensive work to explore the information needs and preferences of members of the public concerning treatment options for depression and anxiety [15,16]. Consumer preference research [15] indicates that some segments of the public are interested in information available on the Internet accessed independently, whereas others are interested in traditional paper format accessed in consultation with health professionals. The IDA provides information in either (1) Web format or (2) brochure format, which may be downloaded from the website by consumers or health care providers. The IDA addresses a wide range of topics judged to be important by consumers [16] with information presented in modules focused on particular consumer questions with up-to-date evidence and references. Many of the topics have not previously been addressed in available information on the Internet and in patient brochures. The IDA differs from many others in that it is intended for use at any point in the treatment process and may be accessed by interested parties independently or in consultation with a health professional—before, during, or after a health care encounter. It may be used when considering treatment options, when making changes to treatment, and when considering stopping treatment. The resources are available on the Internet in English and French [17] (see Multimedia Appendix 1), have a Creative Commons copyright, and may be saved and disseminated in either the Web version or in fact sheet (pdf) versions (see Multimedia Appendix 2). Consistent with the recommendation of the IPDAS criteria that frontline clinicians be involved in the evaluation and dissemination of the IDA, the purpose of this study was to obtain the opinions of primary care providers about: (1) current information resources for primary care patients concerning treatment options for depression; (2) their opinion about the Informed Choices Depression IDA as a resource for patients in primary care; and (3) facilitators and barriers for the dissemination of the resources in primary care settings.

Methods

Focus Groups

A total of 5 focus groups were conducted at 2 sites of the Ottawa Hospital Academic Family Health Team and 3 large primary care clinics operated by the regional health authority in Winnipeg. Winnipeg and Ottawa are medium-sized cities with populations in the range of 600,000 to 1,000,000 in Central Canada and Eastern Canada, respectively. Each of these primary care settings serves high-need patients including the elderly, children, pregnant mothers, and lower income residents. In the Canadian system, there is no charge to the patient for a physician visit, and there is some coverage of drug costs (depending on the province) but limited availability of counseling and psychotherapy by providers other than physicians. The primary care clinics where the focus groups were held have greater access to nurses and in-house counselors than most primary care settings in the country. Table 1 summarizes the clinicians participating in each group.

XSL•FO

RenderX

Table 1. Focus group participants.

Site	Family physicians	Nurses or nurse practitioners	Social workers or mental health workers[a]	Other	Years of practice[b]	Male or female
	N	N	N	N	M (Range)	N
Ottawa 1	2	4	1	1 resident	27.1 (14-39)	2/6
Ottawa 2	5			1 resident	23.8 (12-35)	3/4
				1 pharm[c]		
Winnipeg 1	7	3/2	1	1 resident	9 (1-35)	6/8
Winnipeg 2	2	2/1		1 OT[d]	19.5 (1-41)	2/5
				1 clerk		
Winnipeg 3	5	3/1		1 manager	23.1 (7-40)	4/8
				1 OT		
				1 PCA[e]		

[a]Social workers and mental health workers had generally completed masters' level training.

[b]Years of practice includes clinical staff only and not administrative staff.

[c]pharm: pharmacist.

[d]OT: occupational therapist.

[e]PCA: primary care assistant.

Procedure

The focus groups ran for approximately 1 hour and were held at regular team meeting times when possible so as to engage as many practitioners as possible from each setting. They were facilitated by a psychologist experienced in leading focus groups (ie, JB for Ottawa clinics, JW for Winnipeg clinics). The facilitator ensured that all participants had an opportunity to speak. There was also a notetaker present for each focus group, and participants received a $25 gift card for their participation. Seven questions guided the focus group discussion. The first groups of questions targeted: (1) what questions do providers hear from patients about treatments for depression; (2) how do they handle these questions; and (3) how do they provide information to patients. After these introductory questions, focus group participants received an example fact sheet from the IDA, a list of the fact sheets in the overall IDA, and a business card suitable for use with patients to inform them about the Web address for the IDA. Participants were asked: (4) whether and how they might consider using the materials; (5) the best ways to make this information available to patients; (6) times when this material would be especially relevant; and (7) how might this material be disseminated to health care providers.

Data Analysis

Focus groups were digitally recorded and transcripts created. Written facilitator notes taken during the focus groups were later filled in by reviewing the recordings. A qualitative framework analytic approach [18,19] was used. This approach is used for applied qualitative research that is both grounded and inductive in the accounts and observations of the people studied but starts deductively from clearly identified research questions and preset objectives. During the first stage of data analysis, each group's transcript was reviewed several times by 1 member of the research team (JB, JW, and RW), and

preliminary themes were identified. The transcript was then reviewed on a line-by-line basis, and themes were coded. By the final group transcript, no additional themes were identified, and thematic saturation deemed reached. The coded transcripts were reviewed and compared by a second team member for accuracy and agreement on codes or categories and to search for missed concepts or themes. Disagreements were discussed and consensus achieved. Themes were summarized and interpreted in collaboration among the 3 team members. Participant background data were compiled and means computed.

Results

Focus Group Themes

The Key Role of the Primary Care Provider

An early theme was the key role of the provider, including developing and maintaining trust, responsibility for the overall health of the patient, assessment and diagnosis, reviewing treatment options, and normalizing the patient's experience. One practitioner expressed this role clearly when saying, "the fundamental interaction between my patient and myself is predicated on trust…If you do not have that trust, it is very hard to do anything that leads to anything effective." Practitioners expressed concerns about time pressure for tasks such as familiarizing themselves with patient resources, helping patients to access information, and discussing treatment options.

Questions About Treatments for Depression

Practitioners reported fielding many questions from patients. The most common are about side effects from medication (eg, sexual side effects, weight gain). Several providers said they take many additional questions about medications: "Most people have questions around effects of pharmacotherapy—Do I need

it? What is the effect? How do you decide when I need it? How long will it take to work? How long do I have to be on it?" Other concerns are about becoming addicted, results to expect, when to stop, what happens when one stops, and comparing or changing medications.

Practitioners also described high patient interest in alternatives to medications. Patients ask questions about the effectiveness of counseling, light therapy, herbal remedies, exercise, and self-management including books and Web-based resources. Common questions mentioned specific to counseling included where to access services, provider recommendations for specific practitioners, cost, and how to access free or longer term therapy.

Treatment Barriers

A number of practitioners indicated that they direct patients to psychotherapy, mindfulness, healthy living (eg, exercise) and/or provide counseling themselves. Cost and limited access to resources are common barriers discussed, particularly related to nonpharmacological treatments. Other barriers include patient knowledge, limited social support, depressive symptoms, and concerns regarding stigma.

Sources of Information

It was clear that practitioners develop their own style of responding to patients, some relying mainly on discussion, whereas others using resources such as pamphlets, websites, or mobile phone apps. Main sources of information described included Web-based resources, crisis lines, community support groups, and written information. Some prefer to be familiar with resources before making recommendations. This is challenging because time to review information sources is limited.

The most commonly used written materials are local crisis line pamphlets and cards. Next were Web-based resources, especially for younger patients. Specific websites mentioned were the Canadian Mental Health Association, MoodGYM, AnxietyBC, and local websites (eg, hospital, self-help organization, or the clinic website). "I refer everyone to ementalhealth.ca," indicated 1 provider. When using Web-based sources, some mentioned that they will show patients the source, print out the information, and review during the appointment and encourage the patient to return to discuss. For example, 1 provider said, "I will make myself a note to ask them if they went to the site or read the handout."

Although practitioners described somewhat less frequent use of written information, in part because it is difficult to have the needed information available during the patient encounter, some practitioners indicated that they often used written information.

Concerns About Currently Available Information

Concerns were expressed regarding the quality of available information. Specific concerns related to reading level of materials, access to Canadian information, limited quality information on medication, and a general concern with the quality and validity of information were available on the Internet and in the media (eg, Dr Oz). Too much information was described by a number as not being helpful and sometimes causing alarm (eg, long lists of medication side effects), 1

practitioner stating, "too much information can actually dissuade them from getting well."

Providers' Opinions of Informed Choices IDA

Practitioners responded positively to Informed Choices Web-based IDA. The general opinion was that the resources seemed useful, supportive of providers' messages, and accessible to patients in terms of language, format, and length: "It looks like it is easy to navigate—handouts are bullet points, not a big paragraph—not overwhelming." The option of both handouts and Web-based formats was viewed positively in that the resource was perceived to appeal to different patient groups. Practitioners described the resources as relevant at any point in the treatment process. The importance of resources that patients can share with support persons was also highlighted: one practitioner described the IDA as "information to start conversations."

Response to a sample of a business card with the website address and key information was particularly positive. Practitioners liked the color (bright blue), the graphic, and the convenient size. Consistent with views expressed regarding dissemination methods, some reported that they would use the cards, whereas others would not. One expressed liking the cards because patients "can look at what they were interested in versus what we think they are interested in." A number thought that making the cards available in the waiting room would be convenient for patients and might initiate helpful discussions in the meeting with the practitioner.

In considering the range of topics available in the IDA, additional content areas were suggested such as materials on postpartum depression, pregnancy and breastfeeding, youth and depression, co-occurring disorders, substance use and depression, and links to local treatment resources.

Disseminating the IDA

There was consensus that information needs to be organized and available to practitioners electronically; captured by the following: "I am not blaming anybody; it is just too busy…We have to think of electronic solutions." Practitioners expressed the need to give patients "access to a smorgasbord that is appropriate, edited, and maintained." Some emphasized the need for information to be all in one place, easy to access, and for a coordinated clinic-wide dissemination strategy (eg, standardize practice to include in counseling session). One provider said, "I like having a central repository of a bunch of resources so that when I want to share resources, I can go to one place that will have multiple things so I do not have to search around." Another stated "[we are] like jugglers with 10 problems with that patient…so the last thing on my mind is where is the damn card…I want to make sure the patient is not suicidal." Accessibility suggestions included linking the resource to a shared computer network. The primary care settings varied in terms of Internet and printer access in clinical areas. Practitioners suggested a variety of approaches ranging from linking the IDA to the electronic medical record (EMR) system, having a link to the IDA website in bookmarks, or hosting the fact sheets on the clinic computer desktops. Managing patient

resources is a significant challenge, given the broad range of problems managed in primary care.

When used, written information from the Web was viewed as helpful in reinforcing the provider's message. The importance of tailoring information was highlighted, 1 describing use of written information when there is insufficient time "to really engage [patient] in the conversation you would like." Posters on clinic bulletin boards were not reported as useful in general, although a few suggested resources could be kept in the waiting area; some mentioned printing handouts for patients to take home, particularly for patients without access to computers. "They need the information to go home with."

Practitioners suggested a range of outlets for disseminating material on a wider scale, including the Internet, through professional associations, consumer-focused organizations, pharmacies, schools, employee assistance programs, insurance companies, and word of mouth.

Discussion

Principal Findings

Practitioners reported a need for quality information and resources for treatment of depression for both patients and providers. The findings support the use of multiple information channels (Web and print) to maximize dissemination of health information to clinicians and patients [15]. An increased preference (among both providers and patients) was reported for Web-based resources [20]. Nonetheless, the ongoing utility of written information was also stressed, and preference on source and format of information varied, consistent with research focused on patients [15]. Providers strongly voiced the need for treatment information to cover a broad range of questions about treatment options among varied patient populations at the same time as remaining concise. Many of the common questions identified by primary care providers related to topics covered in detail in the IDA. Although it is beyond the scope of an IDA to increase the availability of resources (a concern expressed by many primary care providers), the IDA has information for the person seeking treatment on identifying appropriate resources and managing barriers related to cost and availability of services. Many providers indicated that there are significant challenges in accessing and maintaining appropriate information for patients in the context of currently available resources. The need for tools to be accessible on the Internet and at low to no cost is supported by other research [9].

The specific Web-based IDA reviewed in this study was evaluated positively for use by providers in interactions with patients and for use directly by patients. Consistent with other IDA research [21], this IDA was viewed by clinicians as a helpful tool for communicating information between patients and providers. They found the range of topics that covered helpful and suggested several topic areas that would be important to develop as the materials evolve. They also provided specific suggestions for where to disseminate the IDA to reach

the public. The response to the IDA seemed to be more positive than that described in a recent review [11] on IDAs in general. This may be because patients dealing with problems such as depression see themselves as taking a more active role in treatment and as having more treatment options than for many other health problems [16]. Furthermore, the IDA was designed for independent use by the consumer, so, it puts less demand on the health care provider's time than some other IDA resources.

In considering dissemination in primary care, providers emphasized the importance of making the material easily available during patient encounters. It is clear that IDAs need to be at providers "fingertips" to get used. Given the expanding technological applications used within health, it will be important for future research to identify which apps (eg, Web, mobile phone technology) are most powerful for getting health information to both providers and consumers. A revision of the website (hosting system, not content) is currently underway so that the IDA may be conveniently viewed on the mobile phones that are increasingly used by young adults as a primary source of information. The use of business cards as a method to inform consumers about the availability of the IDA was strongly endorsed by primary care providers. On the clinician side, different settings have widely varying availability of electronic resources, and it will be important to develop approaches to dissemination that consider this reality and that integrate well with the systems that primary care providers use. Developing approaches to integrating this and other IDAs with common EMR systems will be important for successful implementation. It would be helpful in future research to evaluate the dissemination and implementation of the IDA in routine practice in a variety of primary care settings [11] and to consider the impact of the IDA on the experience of patients and providers.

Strengths and Limitations of the Study

Although the study involved primary care providers in 2 regions of Canada serving varied patient populations, there are a number of limitations to consider. The clinics were large, serving urban areas, and tended to have staff members from several disciplines. The opinions of practitioners working in smaller clinics or serving small and remote communities may differ. Although it is possible that participants were influenced to report positive opinions on the IDA as a result of the involvement of one of the IDA developers in this study, this potential bias was mitigated by interviewing across 2 regions and having another researcher (JB) lead the study in one of the regions.

Conclusions

The Informed Choices IDA [17] was viewed positively, and suggestions for dissemination were provided that are consistent with those of the literature. Namely, a mix of dissemination approaches with increasing importance place on electronic dissemination. These findings will be helpful in the planning of dissemination of treatment information for primary care settings.

Acknowledgments

Conflicts of Interest

John Walker is the principal investigator of the group that developed the Informed Choices IDA. This tool is freely available at no cost, and no income is made from use of the tool.

Multimedia Appendix 1

[PDF File (Adobe PDF File), 401KB - jmir_v18i6e153_app1.pdf]

Multimedia Appendix 2

[PDF File (Adobe PDF File), 1MB - jmir_v18i6e153_app2.pdf]

References

1. O'Connor AM, Rostom A, Fiset V, Tetroe J, Entwistle V, Llewellyn-Thomas H, et al. Decision aids for patients facing health treatment or screening decisions: systematic review. BMJ 1999 Sep 18;319(7212):731-734 [FREE Full text] [Medline: 10487995]

2. Barratt A. Evidence Based Medicine and Shared Decision Making: the challenge of getting both evidence and preferences into health care. Patient Educ Couns 2008 Dec;73(3):407-412. [doi: 10.1016/j.pec.2008.07.054] [Medline: 18845414]

3. Drake RE, Cimpean D, Torrey WC. Shared decision making in mental health: prospects for personalized medicine. Dialogues Clin Neurosci 2009;11(4):455-463 [FREE Full text] [Medline: 20135903]

4. Clever SL, Ford DE, Rubenstein LV, Rost KM, Meredith LS, Sherbourne CD, et al. Primary care patients' involvement in decision-making is associated with improvement in depression. Med Care 2006 May;44(5):398-405. [doi: 10.1097/01.mlr.0000208117.15531.da] [Medline: 16641657]

5. Loh A, Leonhart R, Wills CE, Simon D, Härter M. The impact of patient participation on adherence and clinical outcome in primary care of depression. Patient Educ Couns 2007 Jan;65(1):69-78. [doi: 10.1016/j.pec.2006.05.007] [Medline: 17141112]

6. Loh A, Simon D, Wills CE, Kriston L, Niebling W, Härter M. The effects of a shared decision-making intervention in primary care of depression: a cluster-randomized controlled trial. Patient Educ Couns 2007 Aug;67(3):324-332. [doi: 10.1016/j.pec.2007.03.023] [Medline: 17509808]

7. Stacey D, Légaré F, Col N, Bennett C, Barry M, Eden K. Decision aid.: Wiley Decision aids for people facing health treatment or screening decisions (Review) URL: http://decisionaid.ohri.ca/docs/develop/Cochrane_Review.pdf [accessed 2016-06-01] [WebCite Cache ID 6hwZE8x4p]

8. O'Connor AM, Wennberg JE, Legare F, Llewellyn-Thomas HA, Moulton BW, Sepucha KR, et al. Toward the 'tipping point': decision aids and informed patient choice. Health Aff (Millwood) 2007;26(3):716-725 [FREE Full text] [doi: 10.1377/hlthaff.26.3.716] [Medline: 17485749]

9. White G, Caine K, Connelly K, Selove R, Doub T. Designing Consumer Health Technologies for the Treatment of Patients With Depression: A Health Practitioner's Perspective. Interact J Med Res 2014;3(1):e2 [FREE Full text] [doi: 10.2196/ijmr.2368] [Medline: 24413087]

10. van der Krieke Lian, Emerencia AC, Boonstra N, Wunderink L, de JP, Sytema S. A web-based tool to support shared decision making for people with a psychotic disorder: randomized controlled trial and process evaluation. J Med Internet Res 2013;15(10):e216 [FREE Full text] [doi: 10.2196/jmir.2851] [Medline: 24100091]

11. Elwyn G, Scholl I, Tietbohl C, Mann M, Edwards Adrian G K, Clay C, van der Weijden Trudy, et al. "Many miles to go …": a systematic review of the implementation of patient decision support interventions into routine clinical practice. BMC Med Inform Decis Mak 2013;13 Suppl 2:S14 [FREE Full text] [doi: 10.1186/1472-6947-13-S2-S14] [Medline: 24625083]

12. Volk RJ, Llewellyn-Thomas H, Stacey D, Elwyn G. Ten years of the International Patient Decision Aid Standards Collaboration: evolution of the core dimensions for assessing the quality of patient decision aids. BMC Med Inform Decis Mak 2013;13 Suppl 2:S1 [FREE Full text] [doi: 10.1186/1472-6947-13-S2-S1] [Medline: 24624947]

13. Coulter A, Stilwell D, Kryworuchko J, Mullen PD, Ng CJ, van der Weijden Trudy. A systematic development process for patient decision aids. BMC Med Inform Decis Mak 2013;13 Suppl 2:S2 [FREE Full text] [doi: 10.1186/1472-6947-13-S2-S2] [Medline: 24625093]

14. Shultz CG, Jimbo M. Decision Aid Use in Primary Care: An Overview and Theory-Based Framework. Fam Med 2015 Oct;47(9):679-692 [FREE Full text] [Medline: 26473560]

XSL•FO

RenderX

15.   Cunningham CE, Walker JR, Eastwood JD, Westra H, Rimas H, Chen Y, The Mobilizing Minds Research Group. Modeling mental health information preferences during the early adult years: a discrete choice conjoint experiment. J Health Commun 2014 Apr;19(4):413-440 [FREE Full text] [doi: 10.1080/10810730.2013.811324] [Medline: 24266450]
16.   Stewart D, Walker JR, Beatie B, Reynolds K, Hahlweg K, Leonhart M. Postsecondary students? information needs and pathways for help with stress, anxiety, and depression. Canadian Journal of Counselling and Psychotherapy 2014;48(3):356-374.
17.   Informed Choices About Depression. 2016. URL: http://depression.informedchoices.ca/ [accessed 2016-06-01] [WebCite Cache ID 6hzokXBa9]
18.   Krueger RA. Analyzing and reporting focus group results. Thousand Oaks: Sage Publications Inc; 1997.
19.   Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess R, editors. Analyzing qualitative data. New York: Routledge; 1994:173-194.
20.   Hoffman AS, Volk RJ, Saarimaki A, Stirling C, Li LC, Härter M, et al. Delivering patient decision aids on the Internet: definitions, theories, current evidence, and emerging research areas. BMC Med Inform Decis Mak 2013;13 Suppl 2:S13 [FREE Full text] [doi: 10.1186/1472-6947-13-S2-S13] [Medline: 24625064]
21.   Uhler LM, Pérez Figueroa Rafael E, Dickson M, McCullagh L, Kushniruk A, Monkman H, et al. InformedTogether: Usability Evaluation of a Web-Based Decision Aid to Facilitate Shared Advance Care Planning for Severe Chronic Obstructive Pulmonary Disease. JMIR Hum Factors 2015;2(1):e2 [FREE Full text] [doi: 10.2196/humanfactors.3842] [Medline: 27025896]

Abbreviations

EMR: electronic medical record
IDA: information decision aid
IPDAS: International Patient Decision Aid Standards

Original Paper

"The Record is Our Work Tool!"—Physicians' Framing of a Patient Portal in Sweden

Christiane Grünloh[1,2*], MSc Tech.Lic; Åsa Cajander[3*], PhD; Gunilla Myreteg[4*], PhD

[1]School of Computer Science and Communication, KTH Royal Institute of Technology, Stockholm, Sweden

[2]Institute of Informatics, TH Köln - University of Applied Sciences, Gummersbach, Germany

[3]Department of Information Technology, Uppsala University, Uppsala, Sweden

[4]Örebro University School of Business, Örebro University, Örebro, Sweden

[*]all authors contributed equally

Corresponding Author:
Christiane Grünloh, MSc Tech.Lic
Institute of Informatics
TH Köln - University of Applied Sciences
Steinmüllerallee 1
Gummersbach, 51643
Germany
Phone: 49 2261 8196 6238
Fax: 49 2261 8196 6666
Email: christiane.gruenloh@th-koeln.de

Abstract

Background: Uppsala County in Sweden launched an eHealth patient portal in 2012, which allows patients to access their medical records over the Internet. However, the launch of the portal was critically debated in the media. The professionals were strongly skeptical, and one reason was possible negative effects on their work environment. This study hence investigates the assumptions and perspectives of physicians to understand their framing of the patient portal in relation to their work environment.

Objective: The study uses the concept of technological frames to examine how physicians in different specialties make sense of the patient portal in relation to their work environment.

Methods: A total of 12 semistructured interviews were conducted with physicians from different specialties. Interviews were transcribed and translated. A theoretically informed thematic analysis was performed.

Results: The thematic analysis revealed 4 main themes: work tool, process, workload, and control. Physicians perceive medical records as their work tool, written for communication within health care only. Considering effects on work environment, the physicians held a negative attitude and expected changes, which would affect their work processes in a negative way. Especially the fact that patients might read their test results before the physician was seen as possibly harmful for patients and as an interference with their established work practices. They expected the occurrence of misunderstandings and needs for additional explanations, which would consequently increase their workload. Other perceptions were that the portal would increase controlling and monitoring of physicians and increase or create a feeling of mistrust from patients. Regarding benefits for the patients, most of the physicians believe there is only little value in the patient portal and that patients would mostly be worried and misunderstand the information provided.

Conclusions: Supported by the study, we conclude: (1) The transfer of a paper-based health care process where patients read on paper into a digital process challenges current work practices and has consequences for the work environment. Mostly, this is explained by the changing positions between the physicians and the patient: the latter can drive the process, which reduces the physicians' ability to guide the patient. (2) The physicians' experiences were expressed as worries: patients would not understand the content of the record and become unnecessarily anxious from misunderstandings. The concerns are to some extent based on a generalized view of patients, which might disregard those, who already actively participate in health care. This study hence reveals a need to provide physicians with information about the values for patients from using patient portals. (3) A change of work practices may be beneficial to increase patient participation, but such changes should preferably be designed and discussed with physicians. However, the strong resistance from the physicians made this challenging when launching the patient portal.

(J Med Internet Res 2016;18(6):e167)   doi:10.2196/jmir.5705

XSL•FO
RenderX

Introduction

Patient Portals and Aim of this Paper

The prospect of increasing costs in health care due to demographic changes and the increase of chronic diseases motivates politicians and policymakers to support patients to participate actively in their care. Enabling patients easy access to their medical record is one approach to make them active in managing their own health [1]. As previous paper-based medical records have been digitalized already, it may seem as a natural consequence that patients can access them over the Internet, for example, through secure eHealth services such as a patient portal. Especially, because in numerous countries, the right to access one's medical record is constituted by law. However, the introduction of a patient portal is accompanied by major concerns, especially from health care professionals, who stress, for example, that Web-based access might lead to an increased workload and privacy risks [2]. Research has shown that such concerns discourage health care professionals from embracing the technology, and that for a patient portal to reach its full potential, patients and physicians need to see it as a technology that adds value to care [3]. Miller et al conclude furthermore that research is needed to examine ways how portals can be implemented to address providers' concerns [3].

In Sweden, the introduction of a patient portal led to strong and mostly negative reactions from health care professionals, especially physicians and their trade union. This was shown in a Web survey from 2013, where 82% of the 385 physicians strongly disagreed or disagreed, that Web-based access to medical records is a good reform [4]. Moreover, the same survey also showed that the physicians were negative to giving the relatives of patients access to the medical records and were reluctant toward the eHealth service as such [5]. Another exploratory study has shown that physicians have different assumptions and perspectives that affect their use of technology and how they view patient empowerment [6]. However, so far, there are few qualitative studies on medical professionals' views of patient portals.

The aim of this paper was to further the understanding of the perspectives of physicians in Uppsala when it comes to the patient portal and its effects on their work environment.

Background and Theory

The Patient Portal

In Sweden, the patient portal was launched in Uppsala County Council to its 350,000 patients in 2012 as a part of a large European Union project. This launch was the result of 15 years of work in several projects and efforts that included law changes as well as lawsuits, as the development project ran into accusations of violating, for example, the Work Environment Act (1977:1160) [7]. The overarching aim of the patient portal is to contribute to patient empowerment and patient participation. The portal makes it possible for patients to log in on a Web service and read their health care information and test results and to use about 10 eHealth services. These services include, for example, booking appointments, following referrals, and reading a list of names of all health care professionals who have entered the medical record (so-called "log list").

Two aspects of the patient portal are relevant to this study: (1) patients being able to read their records with or without delay and (2) the log list. The first aspect concerns patients' access to their medical record with or without a delay. At the time of the interviews, only the signed medical notes and test results were shown to the patients. The patients could access their medical records either immediately after their physician had signed the note or if unsigned at the earliest after 14 days. This way of showing the information with a delay from when it was originally written is called a "respite". This functionality has been changed in that the 2-week respite has been removed altogether. Thus, in the current version of the system used in Uppsala County, all records are accessible immediately to patients. The patient could choose after log-in, whether all or only the signed records will be at display. Most patients (98%) chose to see also the unsigned records [8], which would be specifically marked to distinguish them from the already signed notes.

The second aspect, the log list, came about due to patient integrity. The electronic medical record (EMR) is accessible to all public medical clinics in the county. The log list was implemented as a service that makes it possible for patients to look at a list of names of the health care providers who have logged on to their medical record. One underlying idea was that patients would easily recognize names of people familiar to them, such as a neighbor or a relative, who do not have legal permission to read the records. Patients already had the possibility to request a list of those who had read their record; this Web service is a means to simplify this process for the patients.

That citizens would have access to a portal with different patient services, including Web-based accessible medical records, was a controversial issue, and the reactions in media were strong with more than 70 posts in newspapers [7]. The concerns of the health care professionals in this media turbulence were mostly that patients would not be able to make use of the information provided and that medical records were intended for use by health care professionals only. However, concerns were also raised regarding the effect of the new eHealth services on the work environment in health care [7].

The Digital Work Environment in Health Care

The problems with information technology (IT) in health care in relation to usability have been described numerous times, for example, [9] and [10]. There is also research describing how the work environment of health care professionals is becoming more stressful due to factors such as irrelevant and unnecessary work [11] and a focus on efficiency and patient-centered care. However, according to a recent survey with staff in residential

age care units in Sweden, higher levels of person-centered care was associated with higher levels of satisfaction with work and care [12].

The digitalization of previous paper-based documents may simplify work task to such an extent that the original purpose of the document may change. For example, the Medical Informatics Committee of the American College of Physicians outlines in a position paper how the clinical documentation process has developed over time and is now used for multiple other purposes than direct care of the patient [13]. This has led to requirements that influenced the format and content of the documentation [13]. In their policy recommendations, Kuhn et al state that the primary purpose of clinical documentation should be to "support patient care and improve clinical outcomes through enhanced communication" and that patient access to progress notes and medical records "may offer a way to improve both patient engagement and quality of care" [13].

Technological Frames

The term technological frames (TF) was coined by Orlikowski and Gash (1994) and concern the "assumptions, expectations, and knowledge" people use to understand the technology in their organization [14]. Technological frames in that sense do not only concern the role and nature of the technology but also its conditions, consequences, and applications. The artifact itself and the contexts of design and use are formative aspects of the TF [14]. The perception of a new technology can also be regarded as a social phenomenon, in that each individual is exposed to the attitudes of others, and these take part of the formation of the individual's attitudes, beliefs, and values [15]. In the present research, the group of physicians is under investigation, based on the idea that this group is not homogenous: depending on their experiences, work environment, type of patients, and so forth, the TFs of physicians may vary [16].

Previous research has found that the idea of domains of TFs can be used to characterize the interpretations made by participants [14]. Orlikowski and Gash identified the 3 frame domains:

Nature of Technology: people's images of the technology and their understanding of its capabilities and functionality.

Technology Strategy: understanding of the motivation or vision behind the adoption decision and its likely value to the organization.

Technology in Use: understanding of how the technology will be used on a day-to-day basis and the likely or actual conditions and consequences associated with such use.

The implementation of a new technology is often accompanied by skepticism and inertia due to different assessments of the value of the technology in use and the effects on the particular work environment. The framework of TFs is supposed to analyze, explain, and anticipate outcomes around the technological change in organizations [14]. Although frames can be facilitating in ambiguous situations in terms of reducing the uncertainty and providing a basis for taking action, they can also be constraining when they "reinforce unreflective reliance on established assumptions and knowledge, distort information to make it fit existing cognitive structures, and inhibit creative problem solving" [14].

In this study, to reach the aim to understand the physicians' perspectives of the portal and its effects on their work environment, the frame domains from previous research were adopted.

Methods

Interview Content and Data Collection

Semistructured interviews were conducted in the summer of 2013, about 6 months after the patient portal was launched. A total of 12 physicians were interviewed by 3 different researchers. All researchers used the same template for the interviews to cover the required areas of interest. The template consisted of 27 questions (see Multimedia Appendix 1) and was developed in cooperation through a number of meetings. All interviews were done face to face except 1, which was carried out by email. The email interview was initially also planned to be conducted face to face but had to be rescheduled 3 times. After the last appointment had to be cancelled as well, the respondent suggested the possibility to answer via email, which was then accepted by the research team.

Participants

The ambition should be to select respondents in a manner that helps the researcher to learn as much as possible and to find representative respondents [17]. To get access to physicians who were willing to take part in an interview proved to be a greater obstacle than was anticipated. Different strategies were applied to find physicians, for example, contacting heads of departments, mailing lists, and so forth, which makes it impossible to indicate the exact number of individuals who were asked to participate. However, all physicians who were willing to take part were interviewed.

The project succeeded in getting a positive response from physicians in 4 different specialties: orthopedics, oncology, emergency medicine, and internal medicine. The characteristics of the interviewed physicians (N=12) can be found in Table 1.

Table 1. An overview of the interviewees (N=12).

Characteristics		n (%) or mean (range)
Specialty, n		
	Orthopedics (Ortho)	5
	Oncology (Onco)	3
	Emergency Medicine (EM)	2
	Internal Medicine (IM)	2
Gender, n (%)		
	Female	5 (42)
	Male	7 (58)
Work experience in years, mean (range)		14 (2-30)

Analysis

All interviews were transcribed, translated, and repeatedly read by all authors. The translation took place before the analysis due to the international composition of the research team. A thematic analysis was conducted [18] by using the Computer-Assisted Qualitative Data Analysis Software (CAQDAS) *Dedoose* [19]. This Web-based software was used with an additional layer of encryption to meet data privacy requirements and by this, it allowed the researchers to analyze and code the data both independently and jointly. After the familiarization with the data, initial codes were generated informed by the theory of TFs, but still allowing for an inductive approach. The first set of codes was: Concerns, Physicians Patient Relationship, Experience e-health Services, Implementation & Deployment, Medical Records Online, Opportunities, Patient Empowerment, Work Environment.

The Dedoose Training Center [19] was used early in the analysis to evaluate the coding agreement between researchers (intercoder reliability) and to engage in early discussions about codes and possible themes. The training session enabled a comparison between the coders. Joint coding sessions were carried out, in which the understanding of the different codes was discussed and differences in coding style were identified (eg, length of excerpt). During the joint coding, some codes and subcodes were added (Deliver Bad News, Authority, Expertise, Process Driver).

After the coding, all excerpts were exported, thoroughly read through repeatedly, and commented on. The excerpts were printed to facilitate collation, clustering, and the development of a thematic map. The clustered extracts were read again for each theme to review the internal homogeneity (cf. [18]). Part of the analysis process was also the iterative development of a thematic map and the description of the themes in writing, which helped to identify the relationship between the different themes. The quotes used in this paper have been slightly edited to be more readable.

Results

The analysis revealed 4 themes, which are: Work Tool, Process, Workload, and Control (Figure 1). Related to the first 3 themes are the physicians' concerns about patients (Figure 1), for example, patients misunderstanding or not comprehending the records; experiencing undue anxiety, and possibly being harmed by this. The concerns about patients are explicitly stated in relation to these 3 themes and thus presented here as a part of the respective theme instead of as a separate theme in itself.

As stated previously, domains of TFs are *nature of technology, technology strategy*, and *technology in use.* The identified themes in this study are instantiations of the domains of TFs, in that work tool refers to the *nature of technology*, process and workload refers to *technology in use*, and control refers to both, *technology in use* and *technology strategy.*

In the following sections, each of the 4 main themes will be described in detail.

Figure 1. Identified themes from interviews with physicians.
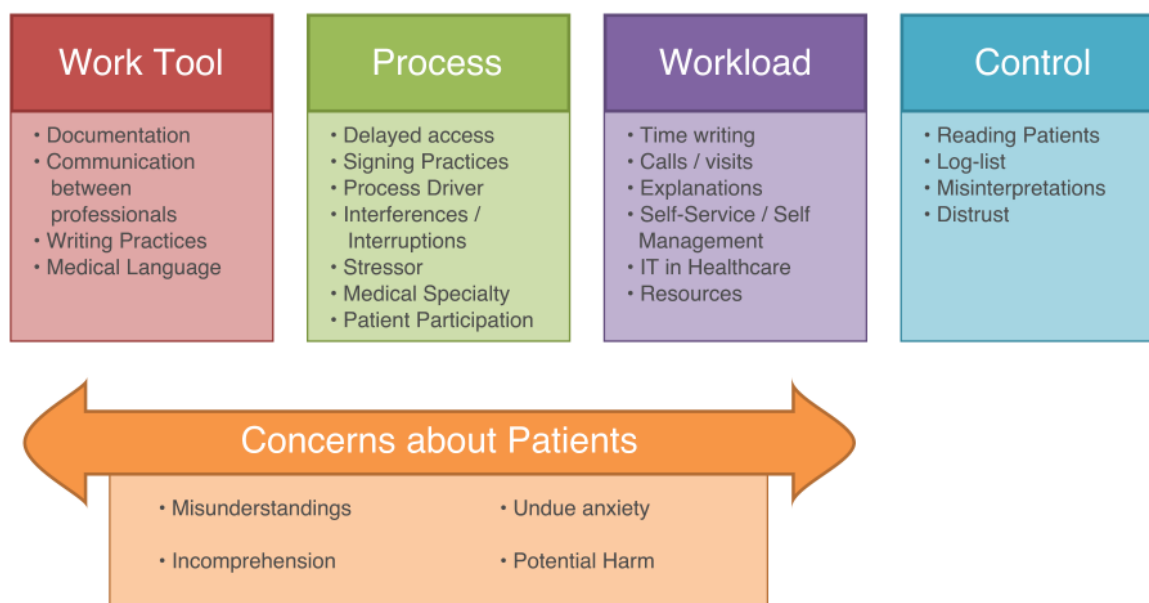
Work Tool

This theme concerns the physicians' view of the EMR as a work tool, used for documentation and communication between health care professionals. Patients reading their medical records through the patient portal raised concerns, especially regarding the difficulties to understand the content due to the use of medical terms.

> *The medical records are written in such a language that I believe many patients have difficulty understanding the text at all.* [Ortho-5]

> *The record is our work tool; it has never been the patients'.* [Ortho-1]

The physicians discussed their work practices in how they write (eg, medical language, suspected diagnosis) and were concerned that the record as a work tool might lose its purpose if one would change the writing to make it more comprehensible for laypeople.

> *If you then are forced to use terms other than the ones you do now, then I do not think you get the same rigor as it has today in the record. Or it may be that many of the information that I might have included if the patient would not be reading the record I will not include now.* [Ortho-4]

> *It is not fair to me as a doctor to change my way of dictating, change my way to express myself for patients to understand. Then the information will be inaccurate, useless, do not fulfill its purpose, and rather be a risk I would say. And it is also a value, /.../ to bring a discussion about tentative diagnoses and stuff that is not perhaps destined for the patient in the first place, /.../ which can also lead to unnecessary anxiety.* [Ortho-2]

Some physicians considered changing their use of the EMR to not writing about suspected serious diseases and instead wait with the documentation until the results are back, and the patient is informed.

> *I am more cautious about how I pronounce that; there could be misunderstanding /.../, they may not understand what it means. /.../ But we await the investigation and then when all the surveys and samples are back, then I would like to meet the patient first, inform, and then write that "now we have progression, the disease is back".* [Onco-1]

Possible changes of writing practices were mainly discussed in relation to the medical jargon or suspected diagnoses and not, for example, in relation to possible offensive notes. Several physicians mentioned, that because the patient always had the right to request a copy, they would already write in a respectful manner. Only few doctors considered adapting their writing slightly, to make use of the system as a way to communicate with the patient directly (eg, include instructions for self-care or different types of reminders for the patient):

> *It is possible that I change a little bit in my way of writing; you can sneak in some messages there. I do not think it is something I avoid writing, but more that, if I know that the patient reads I might add a sentence here and there about things that I want to stress at the end, just for the patient, yes, quit smoking or something like that.* [Onco-2]

However, most physicians did not know, whether their patients read the medical record. Although some of them assumed that this is the case, they do not feel the need to ask the patients about this. One physician assessed that the patient portal had no impact on the medical work:

Yes, it is zero impact to be honest, because no one has been in contact and I feel that I am fairly clear when writing. So if someone has read his medical record then I have not heard about it, since they have in any case not been in touch and said that it is something that they do not understand or something is not right. [Ortho-3]

Process

This theme concerns requirements such as delayed access to and signing of records by physicians, as well as the question, who is the process driver (the patient or the physician). This involves also possible interferences and interruptions by patients, which is related to an increased stress for the physician due to time limitations and challenges on their workflow.

The physicians discussed an aspect that differed requesting paper copies of the record from accessing them over the Internet through the patient portal. To request the paper copy involved a delay that prevented patients to read the record before the physician.

Well, it happens of course very often [that patients request hard copies], but it is seldom that the patient receives any information that you do not already know about /.../ there is a 'delay' anyway, so we have a chance to catch up on the implications of the new findings and also think through a strategy that is also naturally then discussed with the patient at the same time. [Ortho-3]

As described in the Background and Theory section, this delay was built-in into the patient portal by a "respite" when it was launched. Within the first 2 weeks, the patient could only access information that was signed by the physician. Only after a 2-week "respite," unsigned records could be read by patients, which were then specially marked as unsigned. The time in which physicians signed the records differed, and there seems to be no standard practice. Several physicians reported during the interviews that patients had contacted and requested them to sign records, so that the patients can access them.

I got an email from a patient that she wanted to see the test results that I had not signed. I had not signed because it was an elevated test so I wanted to discuss with a colleague and see what we should do. I was told to sign because she wanted to know the answer, and she can not know what it means, she just sees an increased value, and I have to think about other investigations. [Onco-1]

For some physicians, this increased their stress in that not only the records had to be signed on the patient's demand but the physicians also had to "catch up" in terms of collaborating with colleagues and informing patients about results timely.

But now everyone will call, because they have seen the test results and want to know why, so they want a decision at once. So, yes, if all patients would go in and look at the test results, then this is great pressure for us to catch up, have time to write letters, call patients and not to take two-weeks until the next visit, but the day after. [Onco-1]

Some can become calmer actually, when they have seen the test results, and know what they are. But not the oncologic patients, they are mostly worried. [Onco-1]

When the "respite" was removed, patients were able to read the records regardless whether physicians had signed them or not. Hereby patients can read the test results immediately after they have been entered into the record, meaning that a patient might read the results before the physician. This was heavily criticized by the physicians, especially that patients could access the information before they had read it and come to a decision or finished their investigation.

I have a long education, and I do tests and other things and then I put all these things together and in the meantime, while I do the investigation, I do not want anyone to put their nose in this. And then I meet the patient and /.../ I think it can be very dangerous if the patient comes in during this investigation and sees the test results. [Onco-1]

The patient should at least not be able to read before I have signed it! /.../ the final result might differ from the preliminary. It is not good if the patient is allowed to read the preliminary statement. They should only be able to read the final answer, otherwise it is not good. [EM-1]

Patients reading the record before the physician had made sense of it and informed the patient was seen as a risk that could lead to undue anxiety among patients. This anxiety would then lead to patients calling their physician for answers at a time, when the physicians themselves may not have had the time to read up on the results and prepare the talk, or they might get interrupted while dealing with other patients.

The problem is that patients can come in and read information about themselves, and all findings, at a time when physicians often did not have time to read the same results. And even less have been able to discuss the matter and reason with the patient about it, or give notifications at that stage. Many times it requires that you consult with colleagues about the situation before you can come up with a proposal. [Onco-3]

The examples given, related to the situation that patients read something before the physicians, were often rather drastic, for example, patients learning about an initial diagnosis of cancer through reading their medical record at home without a physician present. This example was not only given by physicians specialized in oncology but was indeed common also from other specialties such as orthopedics:

It is also important that we /.../ discuss the meaning of it. What do we do now, and what can this lead to as consequences and all this? That is nothing that you just bring up completely unprepared, but it often requires a lot of reflection. /.../ That the patient is informed about this by reading it at home, and perhaps with an incomprehensible language, it can be terrible anxiety driving. [Onco-3]

I think it is the doctor's task to interpret and filter it, to make it understandable /.../. It could also be that it is a sign of a serious disease, and it should not be told through the Internet as well, alone at home, so I think this is completely wrong. I mean, of course, the patient should have access to their medical records, but not in that way. [Onco-2]

One of the oncologists emphasized repeatedly, that the particular medical specialties in health care have also different requirements:

We are aware of that [the trend of patient empowerment] and we are not opposed to that. But on the other hand, you need to understand that different businesses look different and the conditions, and it is not just about that some are surgeons and others are medicine people and some third are psychiatrists. But it is much more about how we handle the problems that occur and how we make the patient involved in it. And in oncology, we have been taught for decades that this is very important and we spent lots of energy on it, to handle this and then this comes along. Suddenly everything is just thrown overboard, for an entirely new situation that no one has really thought about. [Onco-3]

However, one oncologist identified the opportunity that reading the record could increase patients' participation and encourages them to read, while facing possible consequences:

It increases the participation in healthcare, and many patients can access their test results especially when they know that they will have a planned patient visit. And then I tell them that they should access it, but that they also need to take the consequences that they must wait until their next planned patient visit to discuss the results. And I believe that this suits some patients, but many patients also say that: "I do not dare". And then to just read the notes about the disease contributes to improved engagements. So I am positive to this if it does not result in many problems or harm for the patients. For example, that they get notified at strange times when they cannot contact anyone. But when the patient is aware about this, then I do not see any problems with it. I mean you must take responsibility for your actions and if you want to log in and look for the test results at a certain time, then you have to take responsibility for it even if it is in the middle of the night. [Onco-2]

Workload

This theme concerns physicians' views on the effect of the patient portal on their workload. The aforementioned anticipated need to change writing to be suitable for a layman was 1 aspect that physicians expected to add to their workload. It would not only make the record less efficient for communication with colleagues but would also take more time to write.

In addition, workload was presumed to be affected by requests to change something in the record and an increase of the number of patients' questions. In the interviews, the latter was connected

to assumed lack of medical knowledge by patients, which would increase their demand for explanations. Physicians also expected that when a patient reads, the focus might be put on details (eg, laboratory results) without seeing the overall picture. This could then dominate the discussion, where the physicians would have to explain in more detail why a raised value in this case is not that relevant.

Well, do they understand them then? The sedimentation rate was 23, but the normal is 20, and you have 23, then no one needs to bother about that. /.../ They are not able to interpret the test results, and it leads to more work, revisits and telephone calls, and they are worried. /.../ A thing like this is nothing but a lot of extra work /.../. I see nothing positive in that the patients read their medical records online. [Ortho-1]

Connected to this is also the anticipated need of having to explain, and to convince the patient, what role a specific piece of information should play for the full picture or what treatments are possible alternatives. Physicians envisioned that the patient might use test results to search for alternative treatments in the Internet that might not be appropriate for the specific patient's case or is not a treatment that is used in Sweden.

One can speed up the investigations and that is for the patient's sake, and they can be prepared for what they will hear at the medical meeting. So it may be a good thing with it that they are prepared. They have seen the test results. They know roughly what it could mean and focus on the opportunities available as treatments. They also read about what options they have before they come here; they are well-read in that case. And then there will be greater demands when they come. And it is not certain that they have understood what they have read, but they just saw something and are stuck on what they think is best for them /.../ so it takes time to explain why we do not choose such a treatment. [Ortho-1]

However, some physicians pointed out, that today patients do not understand everything in conversations with the medical professionals either and viewed the patient portal as a good opportunity:

The opportunity is that they are informed in another way and that is good in itself. Because I do not believe that patients understand all times what one says and they can then read it in the record so that they understand. Then it is good. [IM-2]

The opportunity is that they are informed in another way and that is good in itself. Because I do not believe that patients understand all times what one says and they can then read it in the record so that they understand. Then it is good. [EM-1]

I think we are quite bad at giving patients /.../ recommended treatment and plans and those type of things for the future. /.../ Because we say a lot of things /.../ and you know that when the patient returns home he or she will have forgotten 75% of what we have told them. But if we have put it into the record

and patients can go into it and read what we actually said, this must be a good thing. [IM-2]

Some physicians mentioned that patients are better read up today and perceived it as rather positive, that patients may use the record to look for further information in libraries or the Internet to learn more about their condition.

> *Patients can be very informed, and they come with the papers and they can be very well prepared ... and with relatives, too, so then you discuss different things.* [Onco-3]

Some physicians discussed the possibility that workload could remain unchanged or decrease because of patient's possibility for self-service and self-management:

> *Patients have greater insight opportunity and can potentially be given greater responsibility for their own monitoring/follow-up care.* [EM-2]

> *We do not know quite yet, it could generate some phone calls but it could also take away some phone calls so therefore /.../ I do not think there will be an increased burden on the clinic, which is what many of my colleagues are afraid of.* [Onco-2]

However, the physicians related a risk of an increased workload not only to the contact with a patient but also to introduction of IT in health care on a more general level. Here, they transferred their previous (negative) experiences from other IT systems in health care, which had increased their workload:

> *It is very easy that the workload increases, and eventually one becomes more occupied with the systems than with personal meetings. If we take for example, Cytodos and what it takes for a single chemotherapy treatment. It takes a minute to fill out a paper form, but in Cytodos it may take five, ten minutes, so it is a factor of five times, and it can be up to ten to one, at some chemotherapy treatments.* [Onco-2]

In relation to expected risks of increased workload, it was also mentioned in numerous interviews that the work situation is already strained, with a lack of resources and also abounding administrative tasks. Many physicians expressed that there is a need for more staff in health care, to meet and interact with patients.

> *Patients want to e-mail and such things, and I think, it is a great way to communicate. The only problem is that there must be a structure for us /.../ so that it works, so that patients get answers /.../. I think, it is a matter of time and also a question of resources very much. If you want the patients to have more contact with us, get more information, and want us to more directly interact with patients /.../ then we need more people, basically.* [Onco-3]

Control

This theme concerns the physicians' worries that the patient portal might lead to patients monitoring and controlling the physicians. This is related to the possibility for patients to read

their record as such, but often they refer to the log list, which enables patients to see who accessed their medical record.

> *I see no point in this! Is the patient supposed to act as the police?* [EM-1]

Most of the physicians explicitly state that they read records due to professional and medical reasons not for fun or entertainment.

> *The majority of medical personnel who are inside and read a record have no interest in doing it in some kind of aim to crave for sensational news. It is our job to make it as good as possible for the patient and therefore read or write in the record. It is not exciting like reading someone's diary.* [Ortho-2]

The log list has evoked a feeling of distrust in numerous physicians, in that patients are suspicious, mistrust them, and want to guard their information from outsiders. In interviews, it was emphasized that the physicians have a long education and professional experience and that patients should be able to trust them.

> *This is going completely the wrong way! You become a servant from having been the expert who people asked. Now you are a servant that needs to be controlled, because you do not do your best.* [Ortho-1]

> *We speak the truth and now it feels that someone wants to watch us all the time! But we try to do our best! We do not work against patients.* [Onco-1]

In addition, the recorded logs could lead to misinterpretations, for example, when physicians read a patient's record in a consultation in which they might not meet the patient personally (eg, in the role of doctor on duty, or when a colleague asks for a consultation) or if they accidentally logged into a record.

> *You can end up in pretty many records without having done anything wrong, but it can look like you have done something wrong.* [IM-2]

However, some of the physicians expect patients to read their records to ensure that the physician did not write anything by accident, as, for example, including wrong facts. Others saw the log list functionality as a way to reduce the risk of unauthorized access:

> *The advantage is that it is not possible to read someone else's medical records without leaving a trace, which of course hopefully removes any pure curiosity medical record readers, like a neighbor or a relative. I think that is a much less of a problem than people might think.* [Ortho-3]

Discussion

The thematic analysis revealed 4 themes that can be regarded as formative aspects of physicians' TFs. The themes incorporate perceptions of the medical records as their work tool, the repurposing of which would have negative effects on their current processes and increase the workload. Certain aspects of the patient portal are seen as a threat in that it may be used for monitoring the physicians' work. The physicians show concerns

about their patients, who are seen as lacking medical expertise and might get harmed by using the patient portal.

Work Tool

One interpretation of the physicians' view of the record as their work tool is that they see themselves as its owners. They write the content, the information they add to the record is used by them to communicate with other health care professionals, to make a diagnosis, to form decisions, and to select a proper treatment. The transformation of the medical record to a patient portal is seen as time consuming and a threat to the effectiveness of their work tool. Some physicians also expect an upcoming need to change their way of writing, which was experienced as a negative and unnecessary effect on the work environment. In this, they identified possible requirements by the patients reading the notes, which is seen as repurposing the medical record, as discussed also in [13].

The record as such could also be repurposed in that it may serve as a communication tool between physician and patient, which could increase office efficiency as discussed in [3]. However, this transfer was not yet made by most of the physicians, who maintained their current frame of the record as their own work tool. With 1 exception, none of the physicians asked their patients whether they read records nor encouraged them to read. As limited staff engagement is 1 characteristic of low patient portal adoption [20], the lack of encouragement might not contribute to reach the objective of increased patient participation. However, Irizarry et al view the adoption by patients and the endorsement by medical professionals as a natural consequence, *"when existing patient portal features align with patients' and providers' information needs and functionality"* [21].

Process

Before the patient portal, patients could request and read their records on paper. This was not seen as a controversial issue. One interpretation is that the physicians thought of this as a complex and time-consuming process that probably only few patients would undertake for reasons such as insurance claims and so forth. Another interpretation is the aspect of timing. Requesting and accessing the paper copy produces a delay, during which physicians can "catch up": correct records that might have been wrongly dictated or other errors, do all diagnosing, and consult colleagues, and so forth. At the end of this decision-making process, patients were informed by the physician and presented with possible treatment options, and by then, patients would also receive the printout by regular post. The patient portal affected this timing, in that the possibility for a patient to read the record over the Internet removes the previous delay. The roles between patient and physician may potentially be changed.

Physicians refer to themselves as the ones responsible for the caretaking of patients. The physician should make sense of possible test results, make a diagnosis, consult colleagues if needed, and come up with a decision regarding treatment. The interviewed physicians generally preferred to complete all steps before giving information to and having a discussion with the patient. Physicians are trained to deal with medical issues, and

with this mind-set, in addition to their view on patients' difficulties to understand the content in the record, they are concerned that reading the record could harm the patient. In relation to this, only few physicians considered patients with chronic conditions who might be quite knowledgeable and therefore likely be able to understand the content and make use of it. As stated previously, patients reading the record before the physician does, was seen as a risk that could lead to undue anxiety among patients. Anxiety itself was judged to be harmful to the patient. It could also lead to other consequences, such as patients taking the wrong action on their own initiative (ie, ending a medication in advance). The interpretation is that the physicians see the risk of being unable to guide patients and their reactions in exceptional situations.

Web-based access does not only change the process with regard to the information flow, in that patients look up the information before the physician, it might also demand a change that starts well before that. By the time that the tests are taken, the physician might have to inform the patient about possible results, so that patients are "prepared" for what might end up in the medical record. Hence, based on this preparation, patients can decide for themselves, if they want to log-in to the record and accordingly will take the responsibility for it, if the information is bad. This change in the health care process can be seen as taking 1 step closer to patient participation and shared decision making, which is reflected in the phrase "nothing about me, without me" [22]. However, this clashes with the framing most physicians in this study have of their process and the extent to which patients can be included. In addition, a test may reveal an illness that was not anticipated by the physician. In this case, the patient could be caught off guard, although they had been sufficiently informed.

The physicians described, that patients demanding them to sign or asking for explanations at once would interfere with their work processes. A similar concern that patients may expect immediate responses to their requests was also stated in [3]. The physicians in this study perceived the process of gathering all relevant information and preparing themselves before informing the patient as the best way of caring for the patient, preventing undue anxiety. The analysis found "requests to sign" during this part of the process to be seen by the physicians as a negative consequence and an additional stressor for their work environment, especially since they had not been informed about changes made to their work processes. However, physicians might not be aware of the patients' perceptions of this way of processing information. According to Rexhepi et al, the long period of waiting until the physician contacts them and only by this being able to receive the results was reported by patients to be worse than receiving bad results in the Internet by themselves [23].

Most physicians mentioned life-threatening diseases such as cancer when talking about the patient portal, not only the oncologists. The same narrative occurs in numerous interviews and could be seen as a social phenomenon in that the physicians as a group may have collectively constructed or shaped each other's attitudes [15]. One interpretation is that this narrative about patients receiving the initial cancer diagnosis could be a repetition and reproduction of what the doctors read about in

the media at the time of the study. Another interpretation could be that the physicians framed this particular group of patients as the most vulnerable. Most physicians viewed the patients as not knowledgeable when it comes to medicine or medical language and therefore expected them to become anxious through reading the hence incomprehensible record. Interestingly, no physician talked about chronically ill patients, who already do a lot of self-management, for example, diabetes [24]. Having a chronic disease often includes regular laboratory tests to be carried out, the results of which the patients often are able to monitor themselves, and where an appointment is only necessary if the results are out of range.

In addition, what is not discussed in these interviews is the possibility that patients can educate themselves regarding their disease, which can be facilitated when they read their records in a timely manner. Only 1 physician acknowledges that some patients know more about their diseases than the physician, especially if that is a rare disease and the physician has not read up before. This framing of stability of patients' health literacy contradicts, for example, the view of health literacy as a key outcome from health education, which is seen as being critical to empowerment [25].

Although other studies have previously shown that health care professionals are concerned that patients might get worried or confused [3], this study could relate the concerns to a specific view of patients as vulnerable subjects, who do not understand the content of the record and may be harmed by reading it. This idea explains the physicians' perceptions of the preferability of the current workflow, where the physician informs the patient.

Workload

Numerous concerns related to a possible increase of workload were also based on the view that patients will not understand the content of the record. The physicians concluded that this would probably lead to: (1) increased phone calls, (2) longer discussions, and (3) a need to change the writing, which would take more time. However, the concerns about increased phone calls might be unfounded. For example, the cancer patients in the study by Rexhepi et al stated that they would wait for the next visit to ask questions instead of calling and demanding explanations at once [23].

The possibility that the patient portal could be a self-service tool, which possibly reduces workload, was not part of the framing of most physicians. This corroborates the results of [3], in which concerns regarding hampered workflow and increased stress outnumbered the view of potential office efficiency. Instead, the physicians in this study transferred negative experiences from other IT systems, which previously had increased their workload. In addition, it has been mentioned that administrative tasks have increased, leaving less time for the contact with patients. These experiences may have influenced their expectations regarding the patient portal.

Control

Trust, as a vital part of the doctor–patient relationship, was an important aspect within this theme. Physicians referred on several occasions to their long education, in which they learned, for example, how to handle difficult cases and situations.

Physicians had difficulty in seeing benefits of patients reading the records, and some physicians felt as if the patients wanted to control them. This is related to the work by Erlingsdottir and Lindholm [7] on professional autonomy of physicians and the encounter with patient portals. Some physicians indeed see patient portals as limiting their autonomy as professionals.

An interpretation of their framing of the patient portal as a surveillance tool is that due to the perceived lack of benefits for patients reading the medical content, what might be left to motivate patients to read is to monitor doctors' activities in terms of: (1) who logged into the record and (2) whether the doctor entered the information correctly. The aforementioned study on cancer patients did show that patients felt more in control but rather in relation to their care and their own understanding of their health condition [23]. Only few patients mentioned an urge to read because of suspected incorrect entries, and those who found inaccuracies did not file a correction [23]. Furthermore, Huvila et al identified a diversity of patients' positions toward reading their medical records and emphasize that it is important to take these into consideration and find flexible solutions instead of using rather simple demographic groups [26]. These positions were interrelated and did not include monitoring health care professionals as such but instead included the aspect of mistrust and the desire to control their health treatment.

In relation to the log list, most physicians felt the need to explain that there is no reason for them to read the records of *their patients* other than out of professional interest. However, the treating physician does not need to explain this, because from a legal perspective, being the treating physician is the justification for reading a medical record in the first place. Still, that patients read the log list was interpreted by the physicians as a possible mistrust toward the treating physician. An interpretation is that the physicians struggled to take the patient's perspective, who might be concerned that potentially all employees in health care in the whole county can read their medical history, including, for example, neighbors, former life partners, and so forth.

In addition to the previous frames, to regard the patient portal as a tool to monitor the physicians' work may increase the difficulty to identify potential benefits the work environment in health care, but rather stimulates resistance.

Limitations

Limitations of this study include interviewing a number of physicians from different medical specialties where respondents were somewhat unevenly distributed among the fields. In interpretative research, it is however not critical to get a fully representative sample: to reach thick descriptions, which the study did, is more important. Although interviews are considered as a viable method to study TFs [14], a limitation of this study is that the analysis is restricted to the data that can be elicited in interview mode because no additional data could be collected (eg, through observations or surveys).

The interviews took place about 6 months after launch of the system, which could be criticized of being too soon as the participants did not yet know for certain how many patients

made use of the patient portal or how it was used. However, we believe that the use of the concept of TFs is particularly useful in the early stage of deployment because the frames strongly influence the choices made regarding the use of the technologies [14]. The TFs may change over time, so results from this study may be useful as an important starting point to examine these changes. In addition, it would be interesting also to learn whether the physicians are today aware about whether their patients are reading their records. At the time of the interviews, most physicians stated that they neither know nor ask their patients, if they do.

Another limitation can be related to the frame domains. The collected material did not hold a sufficient content to make an in-depth analysis of the frame domain *technology strategy*, although particular questions were asked during the interviews. This can be explained by the fact that the physicians were not informed of managerial and political motives and argumentation behind the implementation.

Conclusions

The analysis showed that the TFs of the physicians were significantly constructed based on assumptions and expectations and not experiences. This is not surprising due to the time of the interview taking place only few months after the launch. However, aspects such as patients' rather stable lack of medical knowledge, demanding explanations, and increasing the workload influenced physicians' overall assessment of the usefulness of the patient portal significantly. The constructed TFs of the physicians helped them to make sense of the patient portal by reducing some of the uncertainties. Hereby, the frames provided a basis for assessing its overall usefulness, as discussed in [14].

The transfer of a paper-based healthcare process, where patients read on paper, into a digital process challenges current work practices and has consequences for the work environment. Mostly, this is explained by the changing positions between the physicians and patients: the latter can drive the process, reducing the physicians' ability to guide patients and their reactions in exceptional situations. Several physicians were concerned about having less time to read, consult colleagues, and to reflect.

The physicians' experiences were expressed as worries: patients would not understand the content of the record, and they would become unnecessarily anxious from misunderstandings. However, some of the frames include assumptions and generalizations regarding patients and their abilities (ie, the patient as health illiterate). Albeit being accurate for some groups of patients, it disregards those patients who already actively participate in health care. This risk is in line with a study on medical students by Johnson et al, who conclude that medical teaching "must make students aware of the ingrained stereotypes that likely influence their perceptions of patients and that form barriers to accurate clinical assessment" [27]. Furthermore, a stereotypical view of patients as passive actors who lack medical knowledge and have to be protected can also significantly influence design decisions of health IT [28]. Thus, the reliance on frames that are based on stereotypes might inhibit a change in health care toward patient participation and influence the design of future IT systems. This study hence reveals a need

to provide physicians with information about the values for patients from using patient portals.

Physicians drew attention to different processes and requirements regarding particular diseases and also in respect to different medical specialties, which might be important to take into account. This is in line with the concept of TFs, which includes context as a formative aspect of TFs [14]. However, patients' experiences should be viewed on a more nuanced level. There is more to being a patient, for example, in oncology, than receiving the initial cancer diagnosis. It is also about dealing with a chronic disease over a long period on a day-to-day basis, where some patients are particularly experienced when it comes to their condition [29]. It is important to enable different stakeholders to express their view from their particular context, for example, the physicians in their particular medical specialty. However, from a patient-centered care perspective, another approach could be to shift focus from medical specialties toward patients, who might be treated by physicians from different specialties at the same time.

As the overarching political aim is to increase patient participation, it is crucial to inform and include physicians in the change process. Patient participation is not obtained through an introduction of a patient portal alone. One example of change is that the patient portal made it possible for patients to access information before physicians. This is changing a work practice and turning a current workflow upside down. It is not surprising that physicians do not welcome this imposed change, which is reflected in this study in their way of finding workarounds (eg, not signing the record; not writing suspicions). Hence, the overall objective of patient participation might not pervade the health care process, if it is not fully understood by physicians, and the technology supporting this goal is seen as a threat imposed top down. A change of work practices may be beneficial to increase patient participation, but such changes should preferably be designed and discussed with physicians. However, the strong resistance from the physicians made this challenging when launching the patient portal.

The results of this study may lead to an impression that all physicians shared a negative attitude toward the patient portal. However, although most of the interviewed physicians expressed concerns in one way or another, the strength of the concerns varied from a really strong to neutral to positive attitude. In addition, the questions during the interview explicitly aimed at exploring both positive and negative concerns. Differences in attitude could be explained by or influenced by: (1) length of work experience, (2) medical specialty, (3) gender, (4) physician's personal experiences (eg, being a chronic patient himself or herself), (5) general attitude toward technology, and so forth. However, the present analysis did not reveal any solid patterns in this small sample.

Further research is needed to investigate the extent of substantiation of the expressed concerns and also how patients actually make use of the patient portal, that is, how they deal with inconclusive or incomprehensible information. Having a more comprehensive view on possible benefits for and current use of the portal by patients could help physicians to consolidate or adapt their TFs accordingly. This could also be a way to

identify possible positive effects of the patient portal on the work environment and how medical professionals could profit from its potential. In addition, well-founded concerns should be addressed in the design of patient portals, together with patients' benefits, that need to be strengthened and communicated to the medical profession.

Acknowledgments

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview Template.

[PDF File (Adobe PDF File), 50KB - jmir_v18i6e167_app1.pdf]

References

1. SUSTAINS. Support Users To Access Information and Services URL: http://www.sustainsproject.eu/ [accessed 2016-02-23] [WebCite Cache ID 6fW2ADJ6C]

2. de Lusignan S, Mold F, Sheikh A, Majeed A, Wyatt JC, Quinn T, et al. Patients' online access to their electronic health records and linked online services: a systematic interpretative review. BMJ Open 2014;4(9):e006021 [FREE Full text] [doi: 10.1136/bmjopen-2014-006021] [Medline: 25200561]

3. Miller Jr DP, Latulipe C, Melius KA, Quandt SA, Arcury TA. Primary Care Providers' Views of Patient Portals: Interview Study of Perceived Benefits and Consequences. J Med Internet Res 2016;18(1):e8 [FREE Full text] [doi: 10.2196/jmir.4953] [Medline: 26772771]

4. Scandurra I, Jansson A, Forsberg-Fransson M, Ålander T. Is 'Patient's Online Access to Health Records' a Good Reform? – Opinions from Swedish Healthcare Professionals Differ. Procedia Computer Science 2015;64:964-968 [FREE Full text] [doi: 10.1016/j.procs.2015.08.614]

5. Ålander T, Scandurra I. Experiences of Healthcare Professionals to the Introduction in Sweden of a Public eHealth Service: Patients' Online Access to their Electronic Health Records. Stud Health Technol Inform 2015;216:153-157. [Medline: 26262029]

6. Cajander Å, Grünloh C, Rexhepi H. Worlds Apart - Doctors' Technological Frames and Online Medical Records. Bamberg: University of Bamberg Press; 2015 Presented at: INTERACT 2015 Adjunct Proceedings. 15th IFIP TC.13 International Conference on Human-Computer Interaction; September 14-18, 2015; Bamberg, Germany URL: https://opus4.kobv.de/opus4-bamberg/frontdoor/index/index/docId/25644 [WebCite Cache ID 6iEy7FJIt]

7. Erlingsdóttir G, Lindholm C. When patient empowerment encounters professional autonomy: The conflict and negotiation process of inscribing an eHealth service. Scandinavian Journal of Public Administration. 2015. (2) p. 27-48 URL: http://ojs.ub.gu.se/ojs/index.php/sjpa/article/view/3142 [WebCite Cache ID 6iEyDAPLc]

8. Lyttkens L. Support Users To Access Information and Services (Grant Agreement No 297206). 2015 Jan 30. Document D6.3 Final Pilot Evaluation Uppsala County Council URL: http://www.sustainsproject.eu/sustainsproject/attachment/d63v12.pdf [accessed 2016-02-22] [WebCite Cache ID 6fUJue03s]

9. Scandurra I, Holgersson J, Lind T, Myreteg G. Development of patient access to electronic health records as a step towards ubiquitous public eHealth. European Journal of ePractice 2013;20:21-36.

10. Statens offentliga utredningar. Effektiv vård. 2016. URL: http://www.sou.gov.se/wp-content/uploads/2016/01/SOU-2016_2_Hela4.pdf [accessed 2016-02-25] [WebCite Cache ID 6fY5Ufxeu]

11. Björk L, Bejerot E, Jacobshagen N, Härenstam A. I shouldn't have to do this: Illegitimate tasks as a stressor in relation to organizational control and resource deficits. Work & Stress: An International Journal of Work, Health & Organisations 2013 Jul;27(3):262-277 [FREE Full text] [doi: 10.1080/02678373.2013.818291]

12. Sjögren K, Lindkvist M, Sandman PO, Zingmark K, Edvardsson D. To what extent is the work environment of staff related to person-centred care? A cross-sectional study of residential aged care. J Clin Nurs 2015 May;24(9-10):1310-1319. [doi: 10.1111/jocn.12734] [Medline: 25420415]

13. Kuhn T, Basch P, Barr M, Yackel T, Medical Informatics Committee of the American College of Physicians. Clinical documentation in the 21st century: executive summary of a policy position paper from the American College of Physicians. Ann Intern Med 2015 Feb 17;162(4):301-303. [doi: 10.7326/M14-2128] [Medline: 25581028]

XSL•FO
RenderX

14. Orlikowski WJ, Gash DC. Technological frames: making sense of information technology in organizations. ACM Trans. Inf. Syst 1994 Apr;12(2):174-207. [doi: 10.1145/196734.196745]

15. Leonardi PM, Barley SR. What's Under Construction Here? Social Action, Materiality, and Power in Constructivist Studies of Technology and Organizing. The Academy of Management Annals 2010 Jan;4:1-51 [FREE Full text] [doi: 10.1080/19416521003654160]

16. Myreteg G. Cost-benefit evaluation of e-health services: acceptance and value creation are interactive forces. Health Syst 2015 Sep 25;4(3):204-211. [doi: 10.1057/hs.2015.10]

17. Merriam SB. Case Study Research in Education: A Qualitative Approach. San Francisco: Jossey-Bass; 1988.

18. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

19. Dedoose. Los Angeles, CA: SocioCultural Research Consultants, LLC; 2015. Version 6.2.17, web application for managing, analyzing, and presenting qualitative and mixed method research data URL: http://www.dedoose.com/ [WebCite Cache ID 6fN5pZJhq]

20. Wald JS. Variations in patient portal adoption in four primary care practices. AMIA Annu Symp Proc 2010;2010:837-841 [FREE Full text] [Medline: 21347096]

21. Irizarry T, DeVito Dabbs A, Curran CR. Patient Portals and Patient Engagement: A State of the Science Review. J Med Internet Res 2015;17(6):e148 [FREE Full text] [doi: 10.2196/jmir.4255] [Medline: 26104044]

22. Coulter A, Collins A. Making shared decision-making a reality. No decision about me, without me. London: King's Fund; 2011. URL: http://www.kingsfund.org.uk/sites/files/kf/Making-shared-decision-making-a-reality-paper-Angela-Coulter-Alf-Collins-July-2011_0.pdf [WebCite Cache ID 6fZO4KuLZ]

23. Rexhepi H, Åhlfeldt RM, Cajander Å, Huvila I. Cancer Patients' Attitudes and Experiences of Online Medical Records. 2015 Presented at: 17th International Symposium on Health Information Management Research (ISHIMR); June 25-26, 2015; University of Sheffield; York, UK p. 19-25.

24. Anderson RM, Funnell MM. Patient empowerment: myths and misconceptions. Patient Educ Couns 2010 Jun;79(3):277-282 [FREE Full text] [doi: 10.1016/j.pec.2009.07.025] [Medline: 19682830]

25. 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]

26. Huvila I, Cajander Å, Daniels M, Åhlfeldt RM. Patients' perceptions of their medical records from different subject positions. Journal of the Association for Information Science and Technology 2015;66(12):2456-2470. [doi: 10.1002/asi.23343]

27. Johnson SM, Kurtz ME, Tomlinson T, Howe KR. Students' stereotypes of patients as barriers to clinical decision-making. J Med Educ 1986 Sep;61(9 Pt 1):727-735. [Medline: 3755759]

28. Grünloh C, Walldius Å, Hartmann G, Gulliksen J. Using Online Reviews as Narratives to Evoke Designer's Empathy. In: Abascul J, Barbosa S, Fetter M, Gross T, Palanque P, Winckler M, editors. Human-Computer Interaction – INTERACT 2015, LNCS 9296. Switzerland: Springer International Publishing; 2015:298-315. [doi: 10.1007/978-3-319-22701-6_22]

29. deBronkart D. From patient centred to people powered: autonomy on the rise. BMJ 2015;350:h148. [Medline: 25670184]

Abbreviations

CAQDAS: Computer-Assisted Qualitative Data Analysis Software
EM: Emergency Medicine
EMR: electronic medical record
IM: Internal Medicine
IT: information technology
Onco: Oncology
Ortho: Orthopedics
TF: technological frames

XSL•FO
RenderX

Original Paper

Adoption of a Portal for the Primary Care Management of Pediatric Asthma: A Mixed-Methods Implementation Study

Alexander G Fiks[1,2,3], MD, MSCE; Nathalie DuRivage[1], MPH; Stephanie L Mayne[1], MHS; Stacia Finch[3], MA; Michelle E Ross[2], PhD; Kelli Giacomini[4], BS; Andrew Suh[1], BS; Banita McCarn[3], MEd; Elias Brandt[5], MS; Dean Karavite[1], MSI; Elizabeth W Staton[4,5], MSTC; Laura P Shone[3], DrPH, MSW; Valerie McGoldrick[1], RN; Kathleen Noonan[1], JD; Dorothy Miller[1], JD, MPH; Christoph U Lehmann[6], MD; Wilson D Pace[4,5], MD; Robert W Grundmeier[1,2], MD

[1]The Children's Hospital of Philadelphia, Philadelphia, PA, United States

[2]The University of Pennsylvania, Philadelphia, PA, United States

[3]The American Academy of Pediatrics, Elk Grove Village, IL, United States

[4]University of Colorado Denver, Denver, CO, United States

[5]The American Academy of Family Physicians, Leawood, KS, United States

[6]Vanderbilt University, Nashville, TN, United States

Corresponding Author:
Alexander G Fiks, MD, MSCE
Department of Biomedical and Health Informatics
The Children's Hospital of Philadelphia
3535 Market Street
Rm 1546
Philadelphia, PA, 19104
United States
Phone: 1 267 426 2304
Fax: 1 267 426 0380
Email: fiks@email.chop.edu

Abstract

Background: Patient portals may improve communication between families of children with asthma and their primary care providers and improve outcomes. However, the feasibility of using portals to collect patient-reported outcomes from families and the barriers and facilitators of portal implementation across diverse pediatric primary care settings have not been established.

Objective: We evaluated the feasibility of using a patient portal for pediatric asthma in primary care, its impact on management, and barriers and facilitators of implementation success.

Methods: We conducted a mixed-methods implementation study in 20 practices (11 states). Using the portal, parents of children with asthma aged 6-12 years completed monthly surveys to communicate treatment concerns, treatment goals, symptom control, medication use, and side effects. We used logistic regression to evaluate the association of portal use with child characteristics and changes to asthma management. Ten clinician focus groups and 22 semistructured parent interviews explored barriers and facilitators of use in the context of an evidence-based implementation framework.

Results: We invited 9133 families to enroll and 237 (2.59%) used the portal (range by practice, 0.6%-13.6%). Children of parents or guardians who used the portal were significantly more likely than nonusers to be aged 6-9 years (vs 10-12, P=.02), have mild or moderate/severe persistent asthma (P=.009 and P=.04), have a prescription of a controller medication (P<.001), and have private insurance (P=.002). Portal users with uncontrolled asthma had significantly more medication changes and primary care asthma visits after using the portal relative to the year earlier (increases of 14% and 16%, respectively). Qualitative results revealed the importance of practice organization (coordinated workflows) as well as family (asthma severity) and innovation (facilitated communication and ease of use) characteristics for implementation success.

Conclusions: Although use was associated with higher treatment engagement, our results suggest that achieving widespread portal adoption is unlikely in the short term. Implementation efforts should include workflow redesign and prioritize enrollment of symptomatic children.

XSL•FO
RenderX

Introduction

Patient portals, Web-based health care applications that enable patients to interact and communicate with their health care providers from outside the office [1], offer a resource to improve communication between patients and clinicians between visits. Patient portal use has increased recently [2]; however, adoption has not been rapid [3], and overall rates of sustained use remain low [4]. Recent research suggests that to effectively engage patients as portal users, several barriers may need to be overcome. For organizations, leadership challenges, marketing problems, and limited staff commitment have constrained portal adoption [5]. Studies have also found that patients who are white and have more health problems are more likely to use portals than others [4,6-8]. In addition, portals have not been as widely used in pediatrics as in the adult setting.

Pediatric asthma is an ideal condition for evaluating the feasibility of implementing portals to facilitate the management of chronic disease in practice. More than 7 million children in the United States have asthma [9], the most common pediatric chronic illness. Asthma is associated with lower quality of life [10,11], more missed days of school for children and work for parents [11-13], higher rates of hospitalization, emergency department visits [14], and death [15]. As appointment follow-up varies in pediatric primary care [16,17], and time constraints of office visits limit discussion, portals may facilitate decision making between families at home and primary care practices. However, the feasibility of using portals to collect patient-generated health information and portals' impact on clinical care across diverse pediatric settings has not been established.

This study evaluated the determinants of implementation success for a portal in pediatric primary care to facilitate communication between families and clinicians regarding treatment concerns and goals, asthma symptoms, medication use, and side effects. In a subset of children with poorly controlled asthma, we further assessed the impact of portal use on asthma management, as clinical impact justifies implementation efforts. Finally, we qualitatively evaluated barriers to and facilitators of portal use experienced by families and primary care practices.

Methods

Setting

Twenty primary care practices were enrolled from 2 practice-based research networks: Pediatric Research in Office Settings (PROS) of the American Academy of Pediatrics and the Pediatric Research Consortium (PeRC) of the Children's Hospital of Philadelphia (CHOP). PeRC is a hospital-owned primary care practice-based research network with 31 primary care practices and 231 clinicians in Pennsylvania and New Jersey [18]; PROS includes 728 practices and 1831 clinicians across the United States and Canada. A convenience sample of 9 PROS and 11 PeRC practices was enrolled.

Study Population

Eligible participants included English-speaking parents or guardians (subsequently "parents") of children aged 6-12 years, who received treatment at a participating practice, had an asthma diagnosis at the time of recruitment, and had an office visit during the past 12 months.

Recruitment

To ensure that low-income children were represented, study practices were required to have ≥20% of children insured by Medicaid or the Children's Health Insurance Program. Each practice was contacted by the investigators (AF, SF), invited to participate, and received an in-person or remote presentation of study procedures. After enrollment, rosters of all eligible children were generated from the electronic health record (EHR) to identify families for recruitment. Because of technical differences in the portals between PeRC and PROS, we tailored recruitment by setting. In PeRC, the study team mailed up to 2 letters to all eligible families, inviting them to call the study team to enroll. Parents provided verbal consent over the phone and enrolled in the MyAsthma portal. In PROS, families were mailed letters with a link to the portal website where families could enroll and consent. Telephone recruitment was used for a random sample of 50 families at each practice who did not respond to the letters. These phone calls were completed by the study team in PeRC and by the primary care practice clinicians/staff in PROS. Informational cards were available and posters were on display in participating offices. Enrolled parents received a $10 incentive for using the portal. PROS practices received $1000 for participating in the study, recognizing the additional work required for data extraction from independent practices.

The MyAsthma Portal

The MyAsthma portal was developed and tested at CHOP to facilitate shared decision making and improve asthma outcomes [19,20]. The portal was designed through a user-centered process including interviews and focus groups with 7 parents of children with asthma and 51 clinical team members, including doctors, nurse practitioners, and nurses. Functions of the portal were designed to reflect features families and clinicians prioritized, and iterative usability testing with parents and clinicians refined the portal system [19]. MyAsthma provides educational material; enables sharing of families' treatment concerns, goals, asthma symptoms, medication adherence, and side effects with the primary care clinical team; tracks asthma control over time for families through the portal and clinicians through the EHR; and

provides decision support to both families and clinicians regarding asthma control and side effects. On enrollment, families entered information about their treatment concerns and goals and completed an asthma control survey that assessed symptoms, medication adherence, and side effects. We used a version of the Asthma Control Test [21] that had been modified slightly to allow for parent proxy report of child symptoms.

Subsequently, families were prompted by email each month to complete the asthma control survey. At CHOP, MyAsthma is embedded within an existing patient portal (MyChart, Epic, Verona, WI, USA; Figure 1). In PROS, families interacted with the portal through a Web interface (Figure 2), and decision support was provided on screen to families and via fax to practices based on asthma control survey results.

Figure 1. The MyAsthma Portal-PeRC Practices. In PeRC, MyAsthma was embedded in an existing patient portal (MyChart, Epic, Verona, WI, USA) already implemented by The Children's Hospital of Philadelphia. ©2014 The Children's Hospital of Philadelphia. All Rights Reserved.

MyAsthma

| MyAsthma Home | Asthma Education | Care Plan | Care Team |

Check-in Survey

✔ You are up-to-date!

Next check-in: Oct 1, 2014

Last check-in: Sep 11, 2014

Goals _Edit_

Parent's Goal: not miss work due to child's asthma

Child's Goal: NOT miss school

Concerns _Edit_

Steroids impacting growth
Cost of medication
Impact on life
Time to manage medications

Other: taste of prednisone

Check-in Survey Timeline

■ No concern   ■ Of concern   ■ Major concern

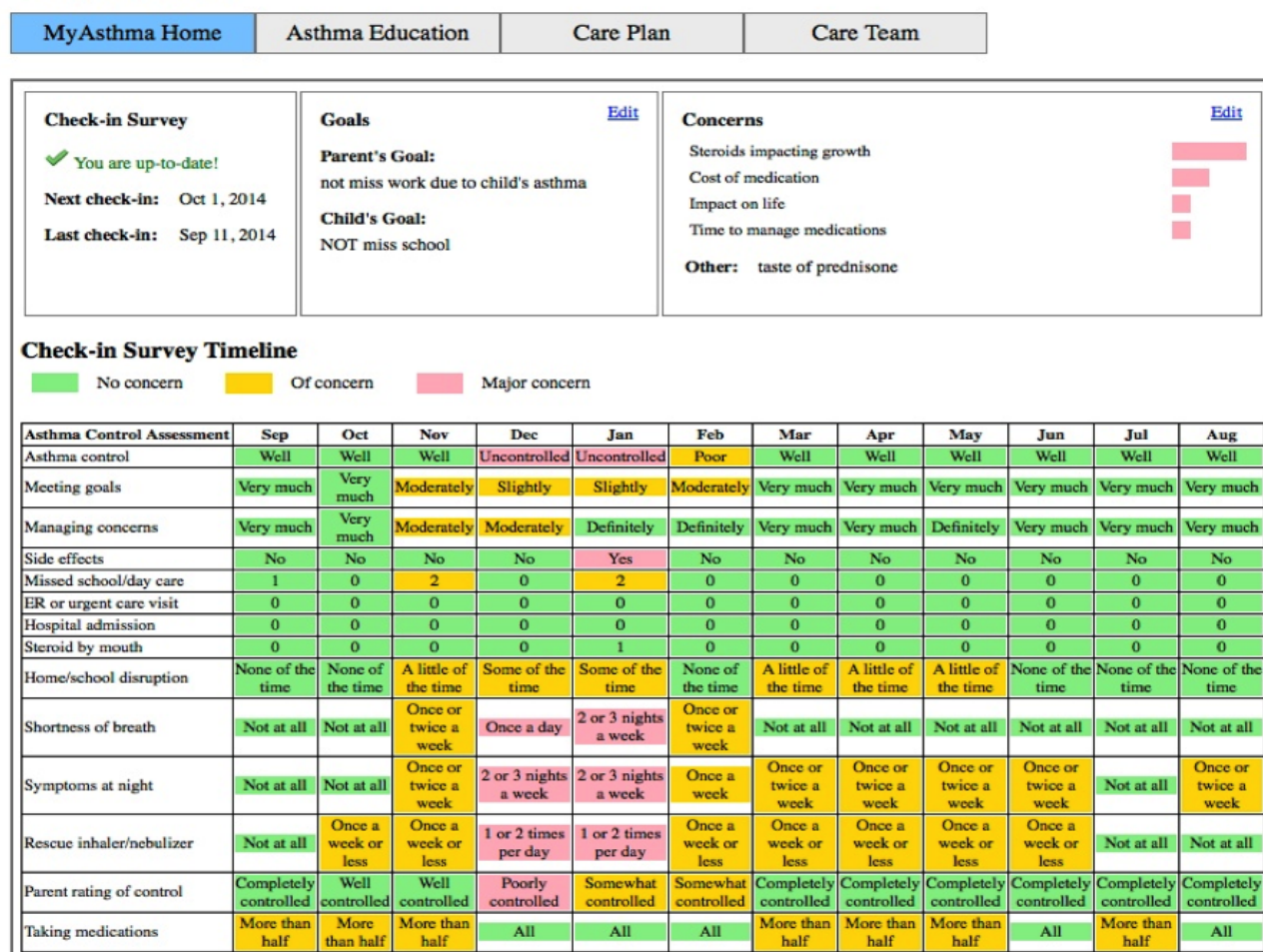| Asthma Control Assessment	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
Asthma control	Well	Well	Well	Uncontrolled	Uncontrolled	Poor	Well	Well	Well	Well	Well	Well
Meeting goals	Very much	Very much	Moderately	Slightly	Slightly	Moderately	Very much	Very much	Very much	Very much	Very much	Very much
Managing concerns	Very much	Very much	Moderately	Moderately	Definitely	Definitely	Very much	Very much	Definitely	Very much	Very much	Very much
Side effects	No	No	No	No	Yes	No	No	No	No	No	No	No
Missed school/day care	1	0	2	0	2	0	0	0	0	0	0	0
ER or urgent care visit	0	0	0	0	0	0	0	0	0	0	0	0
Hospital admission	0	0	0	0	0	0	0	0	0	0	0	0
Steroid by mouth	0	0	0	0	1	0	0	0	0	0	0	0
Home/school disruption	None of the time	None of the time	A little of the time	Some of the time	Some of the time	None of the time	A little of the time	A little of the time	A little of the time	None of the time	None of the time	None of the time
Shortness of breath	Not at all	Not at all	Once or twice a week	Once a day	2 or 3 nights a week	Once or twice a week	Not at all	Not at all	Not at all	Not at all	Not at all	Not at all
Symptoms at night	Not at all	Not at all	Once or twice a week	2 or 3 nights a week	2 or 3 nights a week	Once a week	Once or twice a week	Once or twice a week	Once or twice a week	Once or twice a week	Not at all	Once or twice a week
Rescue inhaler/nebulizer	Not at all	Once a week or less	Once a week or less	1 or 2 times per day	1 or 2 times per day	Once a week or less	Once a week or less	Once a week or less	Once a week or less	Once a week or less	Not at all	Not at all
Parent rating of control	Completely controlled	Well controlled	Well controlled	Poorly controlled	Somewhat controlled	Somewhat controlled	Completely controlled	Completely controlled	Completely controlled	Completely controlled	Completely controlled	Completely controlled
Taking medications	More than half	More than half	More than half	All	All	All	More than half	More than half	More than half	All	More than half	All

Figure 2. The MyAsthma Portal-PROS Practices. In PROS, MyAsthma was available to families through Integrated Health Connect (IHealth Connect), a website developed by the University of Colorado. A test patient is shown.

Outcomes

The primary outcomes included adoption (completion of at least one portal survey during the study period) and sustained use (completion of at least two surveys) of MyAsthma; outcomes were informed by an evidence-based conceptual model of factors influencing implementation success (Figure 3) [22]. We assessed additional outcomes (asthma office visit or asthma medication refill/change within 30 days of survey completion) in a subgroup of children who had uncontrolled asthma according to the results of their first asthma control survey. We focused on these actions because they are appropriate measures to take in response to poor asthma control. These data were extracted from each child's EHR. In addition, parents and guardians reported whether they were more or less likely to (1) speak to their child's doctor, (2) make a change to their child's medication dosage, or (3) make a change to their home environment after using the portal using a 5-point Likert scale.

XSL•FO
RenderX

Figure 3. Conceptual model of factors affecting the implementation of health innovations, adapted from [22].
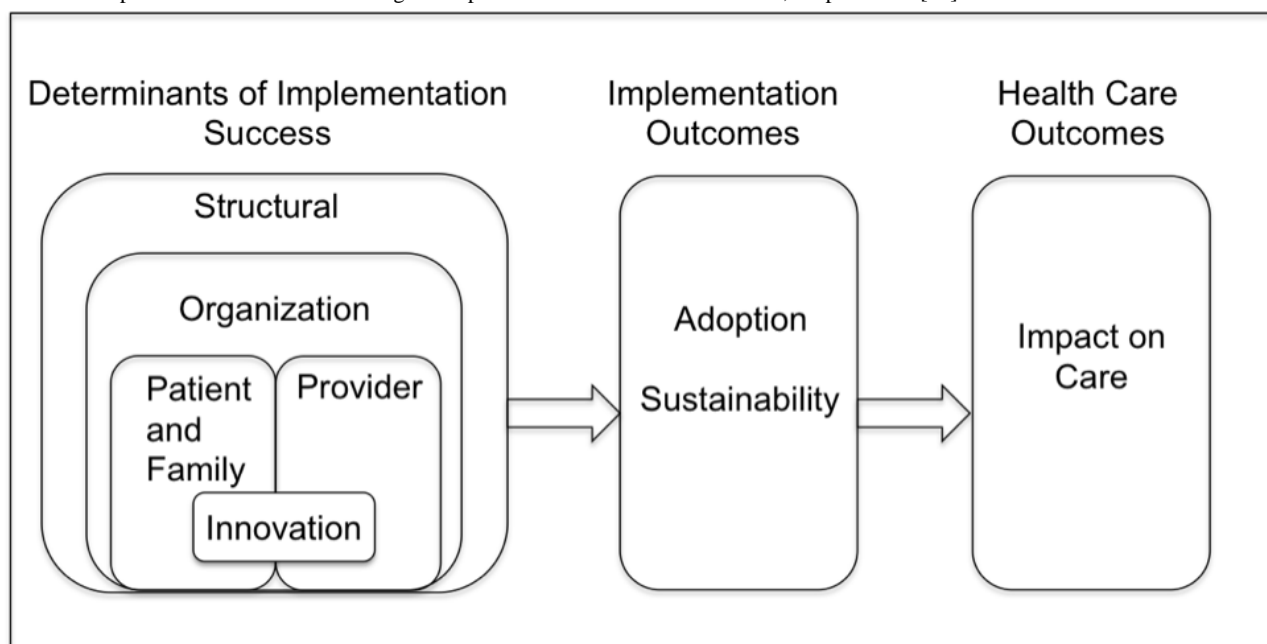
Covariates

We extracted the following covariates from the EHR: patient age, sex, race and ethnicity, asthma severity (mild intermittent, mild persistent, moderate or severe persistent), insurance status (public vs private), and asthma controller medication use at study start (including inhaled steroids, montelukast, combination of inhaled steroid or long-acting β-agonists, and oral steroids). Parent-level covariates were collected via survey from enrolled participants and included age, race and ethnicity, educational attainment, employment status, and relationship to the child. Practice-level covariates included urbanicity (rural, suburban, or urban) and US census region (Northeast, South, Midwest, West).

Statistical Analysis

The study population was described using proportions, means, and standard deviations. Characteristics of children whose parents/guardians completed the portal survey were compared with those of children whose parents and/or guardians did not, using chi-square and t tests. Fisher exact tests were used for categorical data with sparse cell counts, and Mann–Whitney U tests were used for skewed continuous variables. Characteristics of children with sustained use were compared with those of children whose parents or guardians only completed the portal survey once. Multivariable logistic regression was used to model the association of patient characteristics and practice site with portal adoption to identify factors associated with adoption. The proportion of families who enrolled in the portal in response to a mailed letter versus a telephone call was also compared descriptively.

In the subgroup of patients with uncontrolled asthma, we described the proportion of children with an asthma office visit or medication refill or change within 30 days of survey completion. In a sensitivity analysis, we repeated these analyses with a period of 14 days. Furthermore, for each child, we compared these results to the same 1-month period a year earlier to assess whether rates of office visits and medication adjustments changed. We calculated 95% CIs around the change in proportions between years using logistic regression with the margins command in Stata (StataCorp, College Station, TX, USA). We also described parent responses to survey questions regarding the impact of portal use.

All analyses were completed using Stata, version 13.1. The Institutional Review Boards at the American Academy of Pediatrics (reference number: 13 FI 01) and CHOP (reference number: 13-010285) approved this study. All parents provided informed consent and child assent was waived as all information was collected from parents only and because children would not necessarily be readily available when parents were consented by telephone.

Qualitative Study

To evaluate implementation success and identify barriers to and facilitators of portal adoption (Figure 3) [22], trained research assistants on the study team used an interview guide based on our conceptual model to conduct 22 semistructured interviews by phone with parents, purposively sampled to include enrolled (14) and unenrolled (8) from both PROS (7) and PeRC (15), and 10 focus groups (PeRC in-person, PROS by phone, purposively sampled to include diverse representation from both networks) with 46 clinicians. All interviews were recorded then transcribed and coded using NVivo10 (QSR, Cambridge, MA, USA) and interpreted in the context of the conceptual model. Differences in coding were resolved by team consensus.

Results

Adoption and Sustained Use

Few invited families adopted the portal. Out of 9133 eligible patients, 237 (2.59%) completed the portal asthma control survey at least once (adoption). A total of 156 (65.8% of portal adopters, 1.71% of eligible parents) completed the portal survey

more than once (sustained use). Adoption varied widely across practices (0.6%-13.6%; Figure 4). Similarly, sustained use ranged from 0.0% to 13.6%. Reflecting a high level of quality of care, 93.42% of children at PeRC practices with persistent asthma were on a controller medication at baseline. Data on asthma severity were not available in PROS.

Portal users were more likely to have children aged 6-9 years (*P*=.009), to be white (*P*<.001), to be privately insured (*P*<.001), to have mild persistent or moderate or severe persistent asthma (*P*=.002), to be on an asthma controller medication (*P*<.001), and to be receiving a greater number of asthma medications at baseline on average than those who did not use the portal (*P*<.001; Table 1). In addition, those with persistent asthma were twice as likely to use the portal versus those with intermittent asthma (2.37% vs 1.25% at CHOP practices where these data were available, *P*<.001). Sustained portal users were more likely than one-time users to have children who were Hispanic (*P*=.02), have private insurance (*P*=.02), and be from the Northeast (Table 2, *P*=.001). Parents who had sustained use of the portal also had higher educational levels (*P*=.002).

In multivariable logistic regression, the following characteristics were positively associated with portal adoption: receipt of a controller medication at baseline (odds ratio, OR, 2.0, [95% CI 1.5, 2.7]), private insurance (2.0 [1.3, 3.1]), lower child age (1.4 [1.1, 1.9]), and greater asthma severity (1.9 [1.2, 3.0] mild and 1.9 [1.0, 3.5] for moderate or severe persistent versus intermittent; Table 3).

Table 1. Characteristics of families of children with asthma who used the MyAsthma portal compared with families who did not—portal adoption (used portal at least once).

Characteristic at study start			Used portal ≥ once, N (%)	Did not use portal, N (%)	P value[a]
Child characteristics					
	N Children		237	8896	
	Age, years				
		6-9	175 (73.8)	5844 (65.7)	.009
		10-12	62 (26.2)	3052 (34.3)	
	Male		136 (57.4)	5168 (58.1)	.8
	Race[b]				
		White	144 (61.5)	3110 (35.2)	<.001
		Black/African American	75 (32.1)	4789 (54.1)	
		Asian	4 (1.7)	194 (2.2)	
		Other race	11 (4.7)	753 (8.5)	
	Hispanic ethnicity		10 (4.3)	534 (6.1)	.3
	Public insurance[c]		41 (34)	4025 (58.7)	<.001
	Asthma severity[c]				
		Intermittent	49 (41.2)	3857 (57.2)	.002
		Mild persistent	51 (42.8)	2007 (29.8)	
		Moderate/severe persistent	19 (16.0)	873 (13.0)	
	On asthma controller medication		162 (68.4)	4890 (55.0)	<.001
	Mean number of asthma medications (SD)		1.6 (1.4)	1.1 (1.4)	<.001
Practice characteristics					
	Practice Setting				
		Urban	64 (27.0)	4592 (51.6)	<.001
		Rural	52 (21.9)	1309 (14.7)	
		Suburban	121 (51.1)	2995 (33.7)	
	Region				
		Northeast	120 (50.6)	7000 (78.7)	<.001
		South	22 (9.3)	373 (4.2)	
		Midwest	67 (28.3)	960 (10.8)	
		West	28 (11.8)	563 (6.3)	
Parent characteristics [d]					
	N parents completing survey		237		
	Mean parent age (SD)		37.7 (5.8)		
	Relation to child: mother		228 (96.2)		
	Race				
		White	148 (62.4)		
		Black/African American	68 (28.7)		
		Asian	4 (1.7)		
		Other race	17 (7.2)		
	Hispanic ethnicity		17 (7.2)		

Characteristic at study start			Used portal ≥ once, N (%)	Did not use portal, N (%)	P value[a]
	Parent education				
		High school or less	34 (14.3)		
		Some college/associates	81 (34.2)		
		Bachelor's or higher	122 (51.5)		
	Parent employment status				
		Working outside the home	157 (66.2)		
		Self-employed	13 (5.5)		
		Working without pay	43 (18.1)		
		Unemployed	24 (10.1)		

[a]P values calculated using the chi-square test, Fisher exact test, t test, and Mann–Whitney U test.

[b]Race was missing for 53 children (0.6%), ethnicity was missing for 154 (1.7%).

[c]Data on insurance type and asthma severity were only available for PeRC patients (7120 or 78.0% of the total).

[d]Parent characteristics were only collected from families who enrolled in the study (N=237 that completed at least one survey). As such, we are unable to compare these parents with the overall population.

XSL•FO

RenderX

Table 2. Characteristics of families of children with asthma who used the MyAsthma portal compared with families who did not—sustained portal use (used portal more than once).

Characteristic at study start			Used portal more than once, N (%)	Used portal one time only, N (%)	P value[a]
Child characteristics					
	N Children		156	81	
	Age, years				
		6-9	115 (73.7)	60 (74.1)	.9
		10-12	41 (26.3)	21 (25.9)	
	Male		85 (54.5)	51 (63.0)	.2
	Race[b]				
		White	89 (57.8)	55 (68.8)	.2
		Black/African American	54 (35.1)	21 (26.3)	
		Asian	4 (2.6)	0 (0.0)	
		Other race	7 (4.5)	4 (5.0)	
	Hispanic ethnicity		10 (6.5)	0 (0.0)	.02
	Public insurance[c]		27 (28.7)	14 (53.9)	.02
	Asthma severity[c]				
		Intermittent	37 (39.8)	12 (46.2)	.5
		Mild persistent	39 (41.9)	12 (46.2)	
		Moderate/severe persistent	17 (18.3)	2 (7.7)	
	On asthma controller medication		110 (70.5)	52 (64.2)	.3
	Mean number of asthma medications (SD)		1.5 (1.4)	1.8 (1.5)	.4
Practice characteristics					
	Practice Setting				
		Urban	44 (28.2)	20 (24.7)	.2
		Rural	29 (18.6)	23 (28.4)	
		Suburban	83 (53.2)	38 (46.9)	
	Region				
		Northeast	94 (60.3)	26 (32.1)	.001
		South	11 (7.1)	11 (13.6)	
		Midwest	35 (22.4)	32 (39.5)	
		West	16 (10.3)	12 (14.8)	
Parent characteristics					
	N Parents completing survey		156	81	
	Mean parent age (SD)		38.1 (5.5)	37.1 (6.5)	.2
	Relation to child: Mother		149 (95.5)	79 (97.5)	.8
	Race				
		White	93 (59.6)	55 (67.9)	.4
		Black/African American	48 (30.8)	20 (29.4)	
		Asian	4 (2.6)	0 (0.0)	
		Other race	11 (7.1)	6 (7.4)	
	Hispanic ethnicity		13 (8.3)	4 (4.9)	.4

XSL•FO
RenderX

Characteristic at study start		Used portal more than once, N (%)	Used portal one time only, N (%)	P value[a]
Parent education				
	High school or less	17 (10.9)	17 (21.0)	.002
	Some college/associates	46 (29.5)	35 (43.2)	
	Bachelor's or higher	93 (59.6)	29 (35.8)	
Parent employment status				
	Working outside the home	111 (71.2)	46 (56.8)	.06
	Self-employed	5 (3.2)	8 (9.9)	
	Working without pay	25 (16.0)	18 (22.2)	
	Unemployed	15 (9.6)	9 (11.1)	

[a]P values calculated using the chi-square test, Fisher exact test, t test, and Mann–Whitney U test.

[b]Race was missing for 3 children (1.3%), ethnicity was missing for 5 (2.1%).

[c]Data on insurance type and asthma severity were only available for PeRC patients (119 or 50.2% of the total).

Table 3. Child characteristics associated with portal adoption in multivariable logistic regression.[a]

Characteristic at study start		Adoption versus no adoption, odds ratio (95% CI)[a]	P value
Child age 6-9 (vs 10-12) years		1.4 (1.1, 1.9)	.02
Male sex		0.9 (0.7, 1.2)	.6
Race			
	White	Reference	
	Black/African American	0.8 (0.5, 1.3)	.4
	Asian	0.9 (0.3, 2.5)	.8
	Other race	0.5 (0.3, 1.0)	.06
Hispanic ethnicity		0.9 (0.5, 1.8)	.8
Private insurance [b]		2.0 (1.3, 3.1)	.002
Asthma severity [b]			
	Intermittent	Reference	
	Mild persistent	1.9 (1.2, 3.0)	.009
	Moderate/severe persistent	1.9 (1.0, 3.5)	.04
On asthma controller medication		2.0 (1.5, 2.7)	<.001

[a]This model also controlled for primary care practice—odds ratios are not displayed. Practice setting and region were entered into models but dropped due to collinearity.

[b]Data on insurance type and asthma severity were only available for PeRC patients (7120/9133 or 77.96% of the total)—the results presented for these variables are from models including only PeRC participants, whereas the results for all other variables are from models including all participants.

Figure 4. Practice-level variability in portal adoption within 2 pediatric primary care networks. Range 0.6%-13.6%. CIs account for practice size (smaller practices have wider intervals).
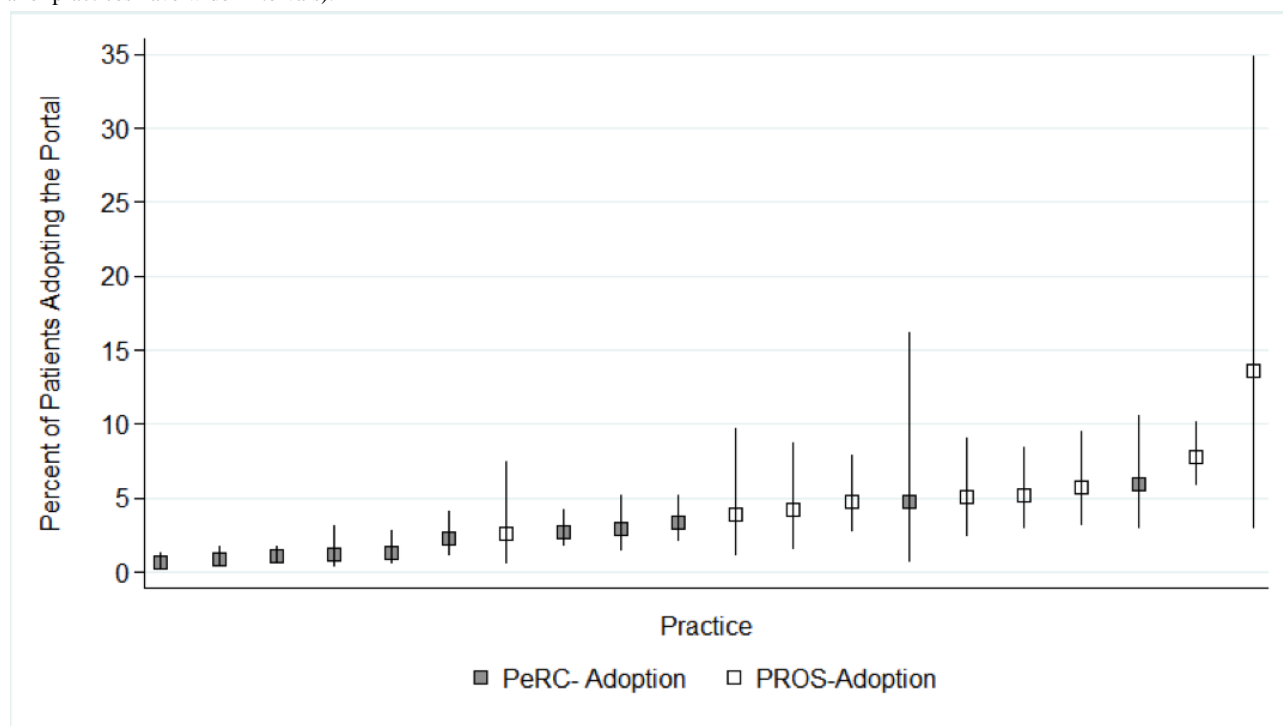
Effect of Phone Versus Letter Recruitment

Portal adopters reported how they learned about the portal. Letters to families resulted in the greatest number of enrolled families. Overall, 208 of 237 enrolled received a letter, 17 received a phone call, 35 heard about it from their child's doctor, nurse practitioner, or nurse, and 3 from an informational card at the practice (25 reported multiple methods). Overall, 2.6% of children contacted by mail only enrolled, whereas 2.7% of those randomized to receive phone calls (they previously received letters) enrolled.

Effect of Portal Use on Asthma Management

Those with uncontrolled asthma commonly planned changes in management after portal use. After completing the first survey, 16% reported an intention to change their child's asthma medication, 27% to contact their child's doctor, and 20% to make a change to their child's environment, with more than one-third (27 parents, 36%) reporting an intention to take at least one action. On follow-up surveys, 22% reported a medication change, 41% reported contacting their child's doctor, and 16% reported making a change to their child's environment (Table 4).

Health records confirmed that portal completion was associated with changes in asthma care. Of the 76 children with uncontrolled asthma after the first survey, 20 (26%) had a medication change or refill within 30 days of survey completion, and 21 (28%) had an asthma-related primary care visit within 30 days (Table 5). These numbers represent a significant increase in medication changes or refills and asthma-related visits when compared with the same period the year prior for each child (14% increase in medication changes [95% CI, 2%, 27%] and 16% increase in visits [95% CI, 3%, 28%]). Results were similar in a sensitivity analysis that examined the 14-day period after portal use.

Table 4. Changes to asthma management planned and taken by families in response to receiving an uncontrolled result on the MyAsthma survey: based on parent survey.

		N (%)
N children with uncontrolled asthma		76
Actions planned as of first survey (parent-reported) [a]		
	Contact doctor	20 (27)
	Change medications	12 (16)
	Change environment	15 (20)
Actions taken as of second survey (parent reported)		
	N uncontrolled with a follow-up survey completed	49
	Contacted doctor	20 (41)
	Changed medications	11 (22)
	Changed environment	8 (16)

[a]Parent/guardian reported being more likely or much more likely to take these actions after completing the MyAsthma survey

Table 5. Changes to asthma management planned and taken by families in response to receiving an uncontrolled result on the MyAsthma survey: based on electronic health record data

Actions taken, based on electronic health record data[a]	Within 30 days of survey completion, N (%) of children	In comparison period (the same 30-day period 1 year prior), N (%) of children	Difference between study year and previous year, N (%) of children (95% CI)
Medication change	20 (26)	9 (12)	+11 (+14% (2, 27))
Primary care asthma visit	21 (28)	9 (12)	+12 (+16% (3, 28))
Either action	30 (39)	14 (18)	+16 (+21% (7, 35))

[a]The denominator for all percentages from the electronic health record-based data is 76 (all children with an uncontrolled result on the first survey)

Qualitative Results

Qualitative results revealed the importance of practice organization, family, and innovation characteristics to portal adoption (Table 6). Few health system factors were discussed and, when mentioned, clinicians disagreed about the value of incentives to promote portal adoption. Only 1 parent interviewed mentioned that an incentive (in this case from the research team) encouraged her to try the portal. For practices or clinicians, 3 primary themes emerged: the need for well-defined and coordinated workflows, the importance of practice responsiveness to portal surveys, and challenges related to identifying children with asthma through the EHR, which resulted in the recruitment of children without recent symptoms. In terms of workflow, clinicians and parents described that portal implementation was facilitated at practices that designated a specific person to coordinate the portal surveys and hampered when workflows were not well defined. Specifically, clinicians in 2 large urban practices reported being short staffed, lacking infrastructure in terms of care coordinators, and uncertainty about the ideal workflow for managing portal surveys. In addition, a perceived need among clinicians for more training diminished enthusiasm at some sites. Among those interviewed, parents of children with well-controlled asthma found MyAsthma less useful if they did enroll. Clinicians, especially in less affluent settings, perceived a lack of computer access as a barrier for parents. At the innovation level, features of MyAsthma that families and clinicians valued included facilitation of communication, increasing family awareness of and responsiveness to uncontrolled asthma, and ease of portal use.

Table 6. Qualitative results of interviews with 22 families and 10 focus groups with primary care clinicians.

Level	Theme	Specific barriers and facilitators and representative quotations
Structural/health system	Financial incentives	*Incentives paid to families may encourage use* (facilitator, 6 practices and 1 enrolled parent)
		"Incentive would grab [parent's] attention. It sounds like [using the portal] doesn't take a lot of time or require a lot of work."—Clinician focus group
		Incentives to families or providers would not encourage adoption (barrier, 3 practices)
		"I don't really know that more money would incentivize…compliance."—Clinician focus group
Practice/clinician	Workflow and coordination	*Coordination of portal surveys by a particular staff member facilitated implementation* (facilitator, 4 practices)
		"… we had a particular person who was spearheading it, so it wasn't like 5 different people were picking up the faxes, they went to a central person and that person distributed it from there, and that I think was helpful. It would've been more confusing would we have had everybody in that."—Clinician focus group
		Lack of an established workflow (barrier, 3 practices)
		"I also did not actively push [portal use] at all. I have fear of MyChart. That I'll have not a good ability to manage the in-basket, and that our support team, while excellent, is already stretched, and not…we haven't built a great infrastructure in terms of care coordinators being able to handle first line, so until we feel secure that's in place and really well running, it feels like we are putting the cart before the horse."—Clinician focus group
		Lack of training for practices impeded effective use (barrier, 6 practices)
		"Triage was not trained and did not know what questions were asked in the portal so found calling patients to follow up difficult." —Clinician focus group
	Practice responsiveness to surveys	*Responsiveness by practices encourages use* (facilitator, 4 practices and 3 enrolled parents)
		"I had a mom that was really happy I called; that the office followed through, she was like I'm really glad you guys called me…it just felt good to type something in and get a response"—Clinician focus group
		"We always received a follow-up phone call from our pediatrician just making sure that we didn't have any questions, so I thought it was a great, you know, communication tactic"—Enrolled parent
		Lack of follow-up by practices discouraged continued use (barrier, 4 enrolled parents)
		"… I didn't really get any feedback or whatever from my doctor either way. So maybe… if it was actually hooked into responses from my doctor, then I would be more apt to use it...there was no interaction with my doctor's office or whatever, with doing that. I didn't really understand...does my doctor, does he see our answers and everything that goes into that."—Enrolled parent
	Identification of children with asthma from the EHR	*Challenges selecting eligible patients using the EHR* (barrier, 2 practices)
		"…the selection process by which the patients were identified, needed, um, tweaking… It was too broad. It identified patients with a diagnosis of asthma in some cases quite a bit distant past. Or they might have had a diagnosis of wheezing per se, not, not asthma. …The family's orientation was 'my child doesn't have asthma.'"
Parent/child	Asthma severity	*Parents of children with well-controlled asthma found less utility in the portal and were less likely to use it* (barrier, 3 practices, 3 enrolled parents, 3 unenrolled parents)
		"I noticed a lot of that was geared towards kids that are pretty severe, having multiple visits, stuff like that. We actually had a pretty mild winter here, so we really didn't have a ton of asthma. We live in a pretty small rural area with pretty clean air so we just don't see the severity that we used to see in [other areas]." —Clinician focus group
		"I guess for someone whose asthma is very well controlled like my son's, it is not really useful. If we were having difficulty then I guess it could have been better but we didn't really need it."—Enrolled parent
		"My son's asthma is not very severe, so I think that if it was a significant daily type of problem for our family then I probably would have been interested in something like that, but we really don't have any trouble at all controlling his asthma. For us, at this point, it is really very simple for us to control. He every once in a while needs his inhaler, and that's about it"—Unenrolled parent
	Computer/Internet access	*Lack of computer/Internet access* (barrier, 5 practices)
		"…There might have been some access issues, we have, definitely a poorer population up here so not everybody has a computer, they might not want to access it on their phone." —Clinician focus group

XSL•FO
RenderX

Level	Theme	Specific barriers and facilitators and representative quotations
Innovation	Communication	*Portal use improved communication between families and primary care practices* (facilitator, 5 practices and 6 enrolled parents)
		"I think functionally the portal was easy, it's a way of patient physician communication to happen without the utilization of an office visit. It was a way of patients checking in saying, this is how its going and so that there's better communication, optimizing the situation and cases where an alert was sent out where the patient really wasn't doing that well. Well, we could move ahead and schedule them and find out why."—Clinician focus group
		" It propelled me to call my doctor more…and to ask the appropriate questions."—Enrolled parent
	Ease of portal sign up and use	*Portal was accessible and easy to use* (facilitator, 1 practice and 8 enrolled parents)
		"I think functionally the portal was very easy."—Clinician focus group
		"It was extremely easy, especially for someone who is not the best on a computer, so it was very straightforward and asked appropriate questions, and easy, honestly, it really was an easy experience…"—Enrolled parent
		"Everything was, it's set up good, it's easy to get on to, if you have questions you wanna ask, I mean, it's simple, it's basic, and I am not a high tech person at all, I can barely use my iPhone without wanting to throw it across the room. So it was actually very easy, very easily accessible."—Enrolled parent
		Time burden involved in completing monthly surveys (barrier, 1 enrolled and 2 unrolled parents)
		I think when I had to report back to the doctor once a week/once a month, I think that's probably what made me, I don't feel like doing that."—Unenrolled parent
	Portal increases family responsiveness to changes in asthma control	*Portal survey increased families' awareness of and responsiveness to changes in asthma control* (facilitator, 2 practices and 7 enrolled parents)
		"At the beginning, I never would have thought that his asthma was uncontrolled, so that was helpful for me to see that his asthma was uncontrolled and now I have it controlled"—Enrolled parent
		"I think it alerted [parents] to signs and send them things to look for. I find that sometimes families are not in tune with what their child's symptoms are. And this kind of alerted them to these are things you need to look for to see if your child is actually under good control. Because unless they are audibly wheezing or going through coughing fits, they wouldn't see it otherwise."—Clinician focus group
		"…After doing the survey seeing where my child was in terms of control, I felt like often times I thought it was well managed but it really wasn't and there were things that I could discuss with the doctor and things that I could do to improve so he didn't go to the emergency room and things like that."—Enrolled parent
		"I may not have been consciously tracking her flares… for me it allowed me to look really in depth about how often was she really having a flare where I may not have been realizing it in the past, and to be better able to track how often her flares are and if she is really controlled"—Enrolled parent

Discussion

Principal Findings

We conducted a mixed-methods, multisite implementation study involving practices from 11 states to assess the feasibility for pediatric primary care practices of using a portal to facilitate communication between clinicians and families regarding asthma treatment, to assess the impact of portal use on asthma care for children with poorly controlled asthma, and to assess barriers and facilitators of portal adoption and sustained use. Overall, we found low rates of portal adoption and sustained use that varied from 0.6% to 13.6% across study practices. However, for those children with uncontrolled asthma, parent use of the portal was associated with a significant increase in asthma medication changes/refills and asthma visits to primary care practices. Qualitative methods underscored the importance of coordinated practice workflows, including practice responsiveness to portal surveys to implementation success.

Parents, especially those with children with uncontrolled asthma, were motivated to continue using the portal because it facilitated a better understanding and tracking of asthma control.

Researchers and health systems in other settings have described low rates of portal adoption. For example, a study of the adoption of a portal for parents of children with cystic fibrosis, juvenile idiopathic arthritis, or diabetes reported that only 28% of invited families obtained a portal account, and only 48% of those (13% total) actually used the portal [23]. Even lower activation rates were observed in a study of a non–disease-specific portal in pediatric primary care, where rates of adoption have been lower than in adult health care settings [24]. Although prior studies found that patients with chronic diseases were more likely than others to register for or use a portal than others [25,26], the adoption rate in our population of children with asthma was quite low. Our qualitative results revealed that, at least in part, the low participation rate resulted from the inclusion of children that

parents perceived had well-controlled asthma. These results are consistent with studies in diabetes that found that patients who believed their disease was well controlled felt that entering information over time was unnecessary [27] and were less likely to enroll [28].

Adoption may also have been limited by practices' infrastructure and workflow for managing electronic receipt of patient-reported information. In our qualitative study, both clinicians and families highlighted the importance of coordinated and responsive workflows to implementation success. Workflow issues have been described previously as a challenge to portal implementation [5,27,29] and a reason for variability in adoption between practices [5,24]. In a case report of the portal adoption experience at 4 different adult primary care practices, practices with strong leadership and high staff engagement had higher rates of enrollment [5]. Learning collaboratives focused on workflow redesign in family medicine practices resulted in rates of portal use exceeding 25% of patients [26]. Especially relevant to the 32 billion dollar Federal Meaningful Use Program in the United States [30], these findings underscore the importance of integrating the portal into office systems and focusing provider and staff attention on their use.

The value to parents, practices, and the health system of implementing portals depends on their ability to improve communication and, ultimately, outcomes. Although adoption of the portal was low, portal use was associated with increased family and practice engagement in asthma management. These results are consistent with our prior pilot trial of MyAsthma, in which clinical outcomes including frequency of asthma flares and days of work missed by parents improved significantly among enrolled families [20]. Our finding that prescriptions and asthma visits increased among uncontrolled patients after using the portal is also consistent with studies among both children with autism and adults with diabetes that showed more active management after enrollment [31,32]. Mechanistically and as supported by our qualitative interviews, portal use may support disease management by improving patient–provider communication [33-36]. These findings support continuing effort to spur portal adoption and sustained use.

Collectively, the results of this study suggest multiple persistent barriers for the use of portals to support chronic disease management in pediatrics and that achieving high rates of adoption likely depends on the extent of existing practice infrastructure focused on disease management. We found that some families invited to participate did not consider their children to have active (or any) asthma. These results highlight the need to cautiously define the population to establish any metric for implementation success. In the case of asthma, researchers have developed definitions of an "asthma computable phenotype," an algorithm based on data such as diagnosis, visits, medication, and laboratories, that accurately identifies patients with asthma [37,38]. Tailoring such definitions will be important for directing portals toward those most likely to benefit.

Limitations

This study had several limitations. First, although we enrolled practices from 11 states and practices varied greatly in adoption, slightly more than half of practices were from a single health system, potentially limiting the generalizability of results. In addition, although the asthma portal was implemented within primary care practices, it was implemented within the context of a research study. Findings may not reflect the results that would be observed if practices implemented a portal themselves. Third, this study had a relatively short follow-up period, limiting our ability to assess sustained use over a longer timeframe. Fourth, in our analysis of changes to medication and visits among children with uncontrolled asthma, we were unable to adjust for asthma severity due to limited sample size. Finally, we focused on a single chronic condition; however, asthma is a common chronic condition for which clinical trial evidence supports improved outcomes with portal use [20].

Conclusion

Despite the potential for real benefits to communication and child health outcomes, results of this multisite implementation study suggest that achieving high levels of portal adoption is unlikely in the short term. Many practices will require redesigned and coordinated workflows and will need to develop targeted outreach to families of children with poor asthma control to ultimately support the use of this technology.

XSL•FO

RenderX

Conflicts of Interest

Dr. Fiks and Dr. Grundmeier are the coinventors of the "Care Assistant" software that was used to implement the portal in the electronic medical record in this study. They hold no patent on the software and have earned no money from this invention. No licensing agreement exists. Dr. Pace has the following conflict of interest: He is the coinventor of the University of Colorado Patient Entered Electronic Recording System (PEERS), which was used as the underlying technology platform for the PROS MyAsthma portal. The remaining authors declare they have no conflicts of interest.

References

1. Yamin CK, Emani S, Williams DH, Lipsitz SR, Karson AS, Wald JS, et al. The digital divide in adoption and use of a personal health record. Arch Intern Med 2011 Mar 28;171(6):568-574. [doi: 10.1001/archinternmed.2011.34] [Medline: 21444847]

2. Silvestre A, Sue VM, Allen JY. If you build it, will they come? The Kaiser Permanente model of online health care. Health Aff (Millwood) 2009;28(2):334-344 [FREE Full text] [doi: 10.1377/hlthaff.28.2.334] [Medline: 19275988]

3. North F, Hanna BK, Crane SJ, Smith SA, Tulledge-Scheitel SM, Stroebel RJ. Patient portal doldrums: does an exam room promotional video during an office visit increase patient portal registrations and portal use? J Am Med Inform Assoc 2011 Dec;18 Suppl 1:i24-i27 [FREE Full text] [doi: 10.1136/amiajnl-2011-000381] [Medline: 21976028]

4. Roblin DW, Houston TK, Allison JJ, Joski PJ, Becker ER. Disparities in use of a personal health record in a managed care organization. J Am Med Inform Assoc 2009;16(5):683-689 [FREE Full text] [doi: 10.1197/jamia.M3169] [Medline: 19567790]

5. Wald JS. Variations in patient portal adoption in four primary care practices. AMIA Annu Symp Proc 2010;2010:837-841 [FREE Full text] [Medline: 21347096]

6. Lyles CR, Harris LT, Jordan L, Grothaus L, Wehnes L, Reid RJ, et al. Patient race/ethnicity and shared medical record use among diabetes patients. Med Care 2012 May;50(5):434-440. [doi: 10.1097/MLR.0b013e318249d81b] [Medline: 22354209]

7. Ralston JD, Rutter CM, Carrell D, Hecht J, Rubanowice D, Simon GE. Patient use of secure electronic messaging within a shared medical record: a cross-sectional study. J Gen Intern Med 2009 Mar;24(3):349-355 [FREE Full text] [doi: 10.1007/s11606-008-0899-z] [Medline: 19137379]

8. Shaw RJ, Ferranti J. Patient-provider internet portals--patient outcomes and use. Comput Inform Nurs 2011 Dec;29(12):714-8; quiz 719. [doi: 10.1097/NCN.0b013e318224b597] [Medline: 21697705]

9. Bloom B, Jones LI, Freeman G. Summary health statistics for U.S. children: National Health Interview Survey, 2012. Vital Health Stat 10 2013 Dec(258):1-81 [FREE Full text] [Medline: 24784481]

10. Guilbert TW, Garris C, Jhingran P, Bonafede M, Tomaszewski KJ, Bonus T, et al. Asthma that is not well-controlled is associated with increased healthcare utilization and decreased quality of life. J Asthma 2011 Mar;48(2):126-132. [doi: 10.3109/02770903.2010.535879] [Medline: 21128880]

11. Schmier JK, Manjunath R, Halpern MT, Jones ML, Thompson K, Diette GB. The impact of inadequately controlled asthma in urban children on quality of life and productivity. Ann Allergy Asthma Immunol 2007 Mar;98(3):245-251. [doi: 10.1016/S1081-1206(10)60713-2] [Medline: 17378255]

12. Dean BB, Calimlim BM, Kindermann SL, Khandker RK, Tinkelman D. The impact of uncontrolled asthma on absenteeism and health-related quality of life. J Asthma 2009 Nov;46(9):861-866. [doi: 10.3109/02770900903184237] [Medline: 19905909]

13. Laforest L, Yin D, Kocevar VS, Pacheco Y, Dickson N, Gormand F, et al. Association between asthma control in children and loss of workdays by caregivers. Ann Allergy Asthma Immunol 2004 Sep;93(3):265-271. [Medline: 15478387]

14. Akinbami L, Centers for Disease ControlPrevention National Center for Health Statistics. The state of childhood asthma, United States, 1980-2005. Adv Data 2006 Dec 12(381):1-24. [Medline: 17180851]

15. Hartman ME, Linde-Zwirble WT, Angus DC, Watson RS. Trends in admissions for pediatric status asthmaticus in New Jersey over a 15-year period. Pediatrics 2010 Oct;126(4):e904-e911. [doi: 10.1542/peds.2009-3239] [Medline: 20876177]

16. Andrews R, Morgan JD, Addy DP, McNeish AS. Understanding non-attendance in outpatient paediatric clinics. Arch Dis Child 1990 Feb;65(2):192-195 [FREE Full text] [Medline: 2317066]

17. Cabana MD, Bruckman D, Bratton SL, Kemper AR, Clark NM. Association between outpatient follow-up and pediatric emergency department asthma visits. J Asthma 2003;40(7):741-749. [Medline: 14626330]

18. Fiks AG, Grundmeier RW, Margolis B, Bell LM, Steffes J, Massey J, et al. Comparative effectiveness research using the electronic medical record: an emerging area of investigation in pediatric primary care. J Pediatr 2012 May;160(5):719-724 [FREE Full text] [doi: 10.1016/j.jpeds.2012.01.039] [Medline: 22364853]

19. Fiks AG, Mayne S, Karavite DJ, DeBartolo E, Grundmeier RW. A shared e-decision support portal for pediatric asthma. J Ambul Care Manage 2014;37(2):120-126 [FREE Full text] [doi: 10.1097/JAC.0000000000000025] [Medline: 24594560]

20. Fiks AG, Mayne SL, Karavite DJ, Suh A, O'Hara R, Localio AR, et al. Parent-reported outcomes of a shared decision-making portal in asthma: a practice-based RCT. Pediatrics 2015 Apr;135(4):e965-e973. [doi: 10.1542/peds.2014-3167] [Medline: 25755233]

21. Schatz M, Sorkness CA, Li JT, Marcus P, Murray JJ, Nathan RA, et al. Asthma Control Test: reliability, validity, and responsiveness in patients not previously followed by asthma specialists. J Allergy Clin Immunol 2006 Mar;117(3):549-556. [doi: 10.1016/j.jaci.2006.01.011] [Medline: 16522452]

22. Chaudoir SR, Dugan AG, Barr CH. Measuring factors affecting implementation of health innovations: a systematic review of structural, organizational, provider, patient, and innovation level measures. Implement Sci 2013;8:22 [FREE Full text] [doi: 10.1186/1748-5908-8-22] [Medline: 23414420]

23. Byczkowski TL, Munafo JK, Britto MT. Variation in use of Internet-based patient portals by parents of children with chronic disease. Arch Pediatr Adolesc Med 2011 May;165(5):405-411. [doi: 10.1001/archpediatrics.2011.55] [Medline: 21536954]

24. Ketterer T, West DW, Sanders VP, Hossain J, Kondo MC, Sharif I. Correlates of patient portal enrollment and activation in primary care pediatrics. Acad Pediatr 2013;13(3):264-271. [doi: 10.1016/j.acap.2013.02.002] [Medline: 23680344]

25. Ancker JS, Barrón Y, Rockoff ML, Hauser D, Pichardo M, Szerencsy A, et al. Use of an electronic patient portal among disadvantaged populations. J Gen Intern Med 2011 Oct;26(10):1117-1123 [FREE Full text] [doi: 10.1007/s11606-011-1749-y] [Medline: 21647748]

26. Krist AH, Woolf SH, Bello GA, Sabo RT, Longo DR, Kashiri P, et al. Engaging primary care patients to use a patient-centered personal health record. Ann Fam Med 2014;12(5):418-426 [FREE Full text] [doi: 10.1370/afm.1691] [Medline: 25354405]

27. Urowitz S, Wiljer D, Dupak K, Kuehner Z, Leonard K, Lovrics E, et al. Improving diabetes management with a patient portal: a qualitative study of diabetes self-management portal. J Med Internet Res 2012;14(6):e158 [FREE Full text] [doi: 10.2196/jmir.2265] [Medline: 23195925]

28. Ronda MC, Dijkhorst-Oei L, Gorter KJ, Beulens JW, Rutten GE. Differences between diabetes patients who are interested or not in the use of a patient Web portal. Diabetes Technol Ther 2013 Jul;15(7):556-563 [FREE Full text] [doi: 10.1089/dia.2013.0023] [Medline: 23777369]

29. Irizarry T, DeVito DA, Curran CR. Patient Portals and Patient Engagement: A State of the Science Review. J Med Internet Res 2015;17(6):e148 [FREE Full text] [doi: 10.2196/jmir.4255] [Medline: 26104044]

30. Centers for Medicare & Medicaid Services. September 2015 EHR Incentive Program Summary Report URL: https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Downloads/September2015_SummaryReport.pdf [accessed 2016-03-04] [WebCite Cache ID 6fl9yGdnr]

31. Bush RA, Stahmer AC, Connelly CD. Exploring perceptions and use of the electronic health record by parents of children with autism spectrum disorder: A qualitative study. Health Informatics J 2015 May 18:NA Epub ahead of print(forthcoming). [doi: 10.1177/1460458215581911] [Medline: 25989804]

32. Grant RW, Wald JS, Schnipper JL, Gandhi TK, Poon EG, Orav EJ, et al. Practice-linked online personal health records for type 2 diabetes mellitus: a randomized controlled trial. Arch Intern Med 2008 Sep 8;168(16):1776-1782 [FREE Full text] [doi: 10.1001/archinte.168.16.1776] [Medline: 18779465]

33. Britto MT, Hesse EA, Kamdar OJ, Munafo JK. Parents' perceptions of a patient portal for managing their child's chronic illness. J Pediatr 2013 Jul;163(1):280-1.e1. [doi: 10.1016/j.jpeds.2013.02.041] [Medline: 23541773]

34. Kittler AF, Carlson GL, Harris C, Lippincott M, Pizziferri L, Volk LA, et al. Primary care physician attitudes towards using a secure web-based portal designed to facilitate electronic communication with patients. Inform Prim Care 2004;12(3):129-138. [Medline: 15606985]

35. Kruse CS, Argueta DA, Lopez L, Nair A. Patient and provider attitudes toward the use of patient portals for the management of chronic disease: a systematic review. J Med Internet Res 2015;17(2):e40 [FREE Full text] [doi: 10.2196/jmir.3703] [Medline: 25707035]

36. Kruse CS, Bolton K, Freriks G. The effect of patient portals on quality outcomes and its implications to meaningful use: a systematic review. J Med Internet Res 2015;17(2):e44 [FREE Full text] [doi: 10.2196/jmir.3171] [Medline: 25669240]

37. Dombkowski KJ, Lamarand K, Dong S, Perng W, Clark SJ. Using Medicaid claims to identify children with asthma. J Public Health Manag Pract 2012;18(3):196-203. [doi: 10.1097/PHH.0b013e31821a3fa7] [Medline: 22473110]

38. To T, Dell S, Dick PT, Cicutto L, Harris JK, MacLusky IB, et al. Case verification of children with asthma in Ontario. Pediatr Allergy Immunol 2006 Feb;17(1):69-76. [doi: 10.1111/j.1399-3038.2005.00346.x] [Medline: 16426258]

Abbreviations

CHOP: The Children's Hospital of Philadelphia

EHR: Electronic health record
PeRC: Pediatric Research Consortium
PROS: Pediatric Research in Office Settings

XSL•FO
RenderX

Original Paper

The Criteria People Use in Relevance Decisions on Health Information: An Analysis of User Eye Movements When Browsing a Health Discussion Forum

Wenjing Pian, PhD, M.Sc. (Information M, BBus(Info Systems); Christopher SG Khoo, PhD; Yun-Ke Chang, PhD

[1]Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore

Corresponding Author:
Wenjing Pian, PhD, M.Sc. (Information M, BBus(Info Systems)
Wee Kim Wee School of Communication and Information
Nanyang Technological University
WKW SCI Building
31 Nanyang Link
Singapore, 637718
Phone: 86 13515020728
Fax: 65 6791 5214
Email: wpian1@e.ntu.edu.sg

Abstract

Background: People are increasingly accessing health-related social media sites, such as health discussion forums, to post and read user-generated health information. It is important to know what criteria people use when deciding the relevance of information found on health social media websites, in different situations.

Objective: The study attempted to identify the relevance criteria that people use when browsing a health discussion forum, in 3 types of use contexts: when seeking information for their own health issue, when seeking for other people's health issue, and when browsing without a particular health issue in mind.

Methods: A total of 58 study participants were self-assigned to 1 of the 3 use contexts or information needs and were asked to browse a health discussion forum, HealthBoards.com. In the analysis, browsing a discussion forum was divided into 2 stages: scanning a set of post surrogates (mainly post titles) in the summary result screen and reading a detailed post content (including comments by other users). An eye tracker system was used to capture participants' eye movement behavior and the text they skim over and focus (ie, fixate) on during browsing. By analyzing the text that people's eyes fixated on, the types of health information used in the relevance judgment were determined. Post-experiment interviews elicited participants' comments on the relevance of the information and criteria used.

Results: It was found that participants seeking health information for their own health issue focused significantly more on the poster's symptoms, personal history of the disease, and description of the disease (P=.01, .001, and .02). Participants seeking for other people's health issue focused significantly more on cause of disease, disease terminology, and description of treatments and procedures (P=.01, .01, and .02). In contrast, participants browsing with no particular issue in mind focused significantly more on general health topics, hot topics, and rare health issues (P=.01, .01, and .01).

Conclusion: Users browsing for their own health issues used mainly case-based relevance criteria to relate the poster's health situation to their own. Participants seeking for others' issues used mostly general knowledge–based criteria, whereas users with no particular issue in mind used general interest– and curiosity-based criteria.

Introduction

People are increasingly seeking for and accessing health information on the Internet. A telephone survey of US adults conducted in 2012 by the Pew Research Center Internet & American Life Project [1] reported that 1778/3014 (59%) of US adults had sought health information on the Internet in the past year. The Pew survey found that among the people who

XSL•FO
RenderX

looked for Web-based health information, 942/1778 (53%) had sought information about specific diseases or problems, 764/1778 (43%) for particular medical treatments, 480/1778 (27%) for weight loss and control information, 338/1778 (19%) for food safety information, and 268/1778 (15%) for drug information. However, it is not clear how people decide what health information is relevant for particular purposes (eg, diagnosis) and in different situations and what relevance criteria they use.

Relevance, relevance judgment, and relevance criterion are difficult concepts to define. It was thought as a vague but well-understood term without a clear definition. In other words, researchers have difficulty defining the term, but users have no trouble deciding whether a document or piece of information is relevant. For the purpose of this study, we define relevance as a user's subjective assessment of the relation between a piece of information or document and the user's situation. A relevant piece of information has a positive relationship to the user's situation, in that it evokes a positive sentiment in the user. In this study, relevance is operationalized in 2 ways:

1. Evaluative relevance: a piece of information or document is deemed relevant if the user says it is relevant.

2. Predictive relevance: a document is deemed likely to be relevant or to contain relevant information if the user selects or clicks on a document surrogate (eg, document title) to view the full document content.

We define a relevance criterion as a reason that contributes to the user's relevance judgment. However, the reason can be expressed at different levels of abstraction. High-level reasons include usefulness, topicality, and quality. This study examined the detailed information in the text as information cues that affect relevance judgment. Hence, in the context of this study, relevance criteria are the pieces of information (expressed in the document) that the reader uses to decide whether the information or document is relevant.

Most of the previous studies on relevance judgment were in the context of information retrieval in Web-based bibliographic and full-text databases [2,3] and more recently in Web-based search engines [4,5]. With the rising popularity of social media websites, it is important to investigate the relevance criteria that people use when browsing social media websites.

There are different types of social media applications with different functional characteristics, resulting in different user behaviors. This study focused on a health discussion forum—HealthBoards [6]. The user posts on discussion forums are organized by topics and subtopics in a hierarchical structure. Users of a discussion forum focus on browsing by topic the posts and responses to posts, rather than checking only the responses to their own posts or the posts of specific people, as in Facebook.

Studies on health information seeking have found that people sometimes purposefully seek health information for their own health issues [7] and at other times encounter useful or interesting information serendipitously [8]. People were found to seek health information in the Internet sometimes for themselves and sometimes for others [1]. This study

distinguished between these 3 types of health information seeking context or needs: (1) seeking information for the user's own health issue, (2) seeking information on behalf of other people (ie, for someone else's health issue), and (3) browsing with no particular health issue in mind. Although systematic review for Web-based health community users suggests that there might be different participation styles among users for different topics [9], this user context or information needs dimension is applicable to any health topic.

Thus, the objective of this study was to find out the relevance criteria that people use to make relevance judgments on a health discussion forum, in these 3 types of use context.

The results of this study carry implications for the design of more user-oriented health information systems. As it was found that people with different types of use contexts focus on different types of health information, the relevance ranking by the search function can assign different weights to different types of information, depending on whether the user is searching for self, others, or with no particular issue in mind.

Prior Work

This section reviews 3 areas of research: studies of relevance criteria, studies of information behavior using eye tracker systems, and factors influencing the Web-based health information behavior.

Studies of Relevance Criteria

Relevance criteria are factors that contribute to users' relevance judgments [10]. They can also be thought of as the clues that people look for to infer relevance [11]. Researchers have investigated users' subjective relevance judgments in various kinds of settings (eg, students, working people, and different occupations) [2,3]. On the basis of their studies, several lists of relevance criteria have been derived.

A list of 21 categories of relevance criteria were derived from an interview of 30 people who used weather information in their jobs about how they judged relevance of information from various kinds of media, such as TV and newspaper. They can be grouped into 10 groups: accuracy, currency, specificity, geographical proximity, reliability, accessibility, verifiability, clarity, dynamism, and presentational quality [3].

Seven groups of evaluation criteria were identified from an interview of 18 students and faculty members to identify how they judged document relevance in preparing assignments [4]: information content, user's previous experience and background, user's belief and preference, other information and sources in the document, the source of the document, the document as physical entity, and the user's situation. Several user aspects were identified in the study, which indicate that the users' own characteristics can influence their relevance judgment, not just the objective features of the information itself.

Further studies were carried out to find out how students made relevance assessments. History-major graduate students were recruited, and they made 27,000 relevance assessments on interview segments of Holocaust survivors and real user topics [12]. Four types of relevance were identified: direct relevance, indirect relevance, context relevance, and comparison relevance.

It was noted that comparison relevance makes use of similarity criteria to help users understand the topic. They divided comparison relevance into external factors (such as time and place), the factors of participants, and factors of act and experiences (such as attitude, feeling, treatment, and experiences of activities). Comparison relevance may be used by users of a health discussion forum to assess the similarity of the poster's health condition (based on the description in the post content) to the user's own condition. In addition, the user's current treatment regime and experiences may be important criteria in his or her relevance judgment.

Other studies have found relevance criteria that are related to the work environment, task, problem situation, and emotional state of the user. One study was conducted to investigate how users judge information relevance in the social Q&A website, Yahoo! Answers [5]. It identified 6 groups of criteria by analyzing the content and attributes of the questions and answers: content, cognitive, utility, information sources, extrinsic and socioemotional, most of which have been identified in previous studies. However, it was found that the socioemotional group included the social aspects of the environment, and the importance of each relevance criterion varied depending on the topic and environment.

These relevance criteria belong to rather broad high-level categories. They can have different meanings in different contexts. No specific health information relevance criteria were proposed in previous studies. This study has derived a list of detailed relevance criteria used in assessing user-contributed health information (the details of the derivation, and the list is shown in Appendix 1).

Information Behavior Research Using Eye Trackers

In the information behavior research area, there have been a few studies that investigated users' eye movements on information objects to understand users' mental processes such as relevance judgment and quality judgment. The duration of eye fixation has often been used as an indicator of user's mental information processing in relevance judgment.

A few researchers have carried out content analyses of information focused on by users [4,13]. Eye movement data were combined with users' post-experiment commentary on their feelings, thoughts, and intentions when performing particular actions on the screen [13]. It was found that eye fixation was associated with cognitive processing, and that a replay of the eye movement recording helped participants to recall the instances of information encountering. In this study, we assumed that when people's eyes focus on a particular text passage, they are interpreting the information conveyed by the passage and making use of it in the relevance decision-making process.

A method was developed to connect eye movement data to users' relevance criteria used during the judgment process when a study was carried out to find out how people make relevance judgments on search results from the search engine, Google [4]. The study participants were asked to think aloud during their search session. The surrogate records in the search result screen with eye fixations were coded by duration and number of fixations and associated with relevance criteria identified in the verbal protocol. They found that eye movements, particularly the attributes of eye fixations (number of fixations and their duration and frequency) reflected users' relevance dynamics. For example, users put less effort on information related with topicality when they examined the search results, as the surrogate records displayed were assumed to be on the topic. However, when deciding that a surrogate record is not relevant, users tended to have longer fixation duration when considering topicality and scope.

This approach of linking qualitative data of users' verbalization with the analysis of eye movements can yield more insights into the relevance decision process. However, when people think aloud, they tend to think differently than in natural situations [14]. People were found to have different reasoning processes than the normal reasoning process when they were required to think aloud. Hence, the think aloud process was replaced in this study with a post-experiment user commentary on what they were thinking during the browsing process.

Factors Influencing the Web-Based Health Information Behavior

Previous studies found that several kinds of factors influence different aspects of consumers' health information behavior. Gender was found to influence the frequency of use of the Internet for health information [15,16]. Women were found to have more health information–seeking behaviors than men. Cultural difference was found as a factor influencing what kind of health information platform people prefer to use [17]. Eastern culture was found to show preference to user-contributed health information such as health social media websites, whereas western culture was found to show preference to expertise-based platforms such as WebMD. Familiarity with the topic was found to influence the search efficacy. If more familiar with the health topic, less modification was made during health information–seeking process [18]. However, there was no study on the particular factor—different types of use context or information needs mentioned previously. This factor exists in most Web-based health information–seeking behavior. Once you start seeking health information in the Internet, you search for your own problem, other's problem, or without a particular problem and for fun. It is important to know how this factor influences people's judgment on the relevance of Web-based health information—a particular kind of health information behavior.

Methods

Framework

This is a study of consumer health information seeking on a particular type of social media website—a health discussion forum. This study focused on the relevance criteria used by people when browsing for health information. People's eye movement behavior of skimming over the text and focusing on particular text passages during browsing was captured with an eye tracker machine. Content analysis of the data was performed to identify the kinds of information people skimmed over and focused on when making relevance judgments, that is, deciding

whether a post contains relevant information to infer the relevance criteria used during the process.

The structure of a discussion forum bears some similarity to information retrieval systems. The user can select a topic by either entering a query keyword in a search box or browsing a classified directory of topics and subtopics. Having selected a topic, the system displays a summary result screen displaying a list of post surrogates (usually the post header, including the title, author, and the number of user views and replies). The user has to scan the post surrogates to select posts that are likely to contain relevant information. When a post surrogate is clicked on, the system displays a detailed post screen showing the content of the main post and response posts from other users. This is sometimes called a discussion thread. In HealthBoards.com, the responses to a post are appended to the end of the post content page, so a post and its responses are equivalent to a full-text document in an information retrieval system. The user reads the post and its responses to identify relevant information. The user also consumes the information by learning something.

The framework used in this study to distinguish between the different stages of health information seeking within a health discussion forum and the 2 types of relevance judgments associated with each stage are summarized in Table 1.

Table 1. Stages of health information seeking within a health discussion forum and the 2 types of relevance judgment associated with each stage.

Stage	User information behavior	Relevance judgment
Stage 1. Searching or browsing		
	Searching or browsing	System retrieves and displays document or post surrogates matching the search query or selected directory category (relevance judgment by machine)
Stage 2. Scanning (the document or post surrogates in the summary result screen)		
	Skimming (over the list of surrogates)	Select surrogates for attention or focus (unconscious relevance judgment based on keywords in the text that catches the eye)
	Examining (individual surrogates)	Select surrogates to retrieve the associated (linked) document (predictive relevance judgment based on an estimation of the likelihood that the document contains relevant information)
Stage 3. Reading (the document or post content)		
	Skimming (through the document or post content quickly)	Select text for attention or focus (unconscious relevance judgment)
	Examining (and absorbing the information in the document or post content)	Deciding whether the information is relevant or likely to be useful in a use context (evaluative relevance judgment)

This framework of user searching and browsing in an information system was derived from 2 frameworks: the 2-stage relevance judgment model [19,20] and the 2-stage browsing model [21,22].

A 2-stage relevance judgment model was proposed for Web-based information searching, based on previous work [23,24]. In this model, the users' overall relevance judgment of a document retrieved by an information system is divided into 2 stages, predictive relevance judgment and evaluative relevance judgment, as detailed in the following section:

Predictive relevance judgment of the document surrogate: the user scans the document surrogates (mainly document titles) retrieved by the information system to make judgments of the likelihood of relevance of the content of the documents. On the basis of the estimated likelihood, the user may retrieve the full text of selected documents to read.

Evaluative relevance judgment of the document content: the user reads some or all the text in the document and makes a relevance judgment based on the information in the text.

These 2 stages reflect the typical design of information retrieval systems that require the user to first assess the list of document surrogates in the summary search result screen for potential relevance and then select (ie, click on) document surrogates that are potentially relevant to read the detailed document content. Predictive relevance judgment is based on an estimation of the likelihood of relevance. This kind of estimation is based on limited information provided by the surrogates, which do not provide enough information for full relevance judgment. On the other hand, evaluative relevance judgment is based on full information in the document. The different nature of these 2 types of relevance judgment may involve different relevance criteria or different types of reasoning.

This 2-stage relevance judgment model is suitable for the study of relevance judgments in health discussion forums because the forum systems typically offer 2 kinds of pages: the summary result page, which lists post surrogates (with post title, number of user replies, number of views, and time of posting) and the detailed post page with full content of a particular post and replies from other users.

The second part of our research framework recognizes 2 types of eye movements—skimming and examining—during information browsing. This skimming or reading model was developed to investigate users' relevance judgments during the information seeking process [21]. It categorized people's reading of retrieved documents into several types of skimming and reading behaviors, according to whether the user's eye gaze moves quickly or is statically focused. It calculated the ratio of cumulative reading to cumulative skimming as a measure to predict the user's relevance decision on a particular document. They found that the ratio of examining to skimming is positively associated with the likelihood of judging information as relevant. This study makes use of the term *examining* instead of the term *reading* used in their study to better reflect that the user is closely focused on particular pieces of text information in contrast to skimming. We use the term *reading* for the entire stage of checking the detailed post content.

Skimming and examining are found not only in the reading stage (ie, reading the document content) but also in the stage of scanning search surrogates. Students were found to exhibit skimming and examining during not only the stage of reading the retrieved full-text documents but also the stage of checking document surrogates to estimate likelihood of relevance [22].

In this study, the 2 types of eye movement behavior, skimming and examining, were incorporated into the 2-stage relevance judgment model, as summarized in Table 1.

Study Setting

This study chose a particular health discussion forum—HealthBoards [6] as the representative. It was chosen from 10 candidates of health forums searched by Google search engine. The details of the 10 candidates are shown in Appendix 2. The HealthBoards.com was chosen for its huge number of users, frequency of use, comprehensiveness, and ranking on rating websites. The details of these reasons are detailed in the following points:

- Number of registered users: 1,079,219 registered users as of September 20, 2015. The second largest was PatientsLikeMe with 300,000 registered users.
- Number of posts: 879,065 threads, and 4,874,692 posts and replies.
- Number of subsections on particular health conditions and problems: more than 280 subsections.
- Number of daily Internet users: 3000+.
- Ranking: No. 1 health forum in Yahoo Health search.

This study recruited research participants from residents who were living in Singapore. Requirements were set to carry out the study successfully:

- Age: 18 to 50 years;
- Education level: either undergraduates or graduates;
- Health condition: no critical disease, including HIV, cancer, and so forth;
- English fluency: competent in English reading and speaking.

The health condition criterion was used to ensure that the participants did not have a debilitating or critical disease because such patients are expected to have rather different needs and information behavior. The English fluency was used to ensure the participants understand clearly the English text on this health forum.

Study Design

The steps of this study were adapted from previous studies using eye tracker machines [13,21] as follows:

Brief the participant on the general purpose of the study.

Ask the participant whether he or she has some health issue to seek information for in the discussion forum. If no, ask whether the participant wants to seek information for the health issue of a relative or friend. If the participant cannot think of a health issue for self or others, then, ask the participant to browse the discussion forum for health information. Ask the participant to describe the health issue in some detail if the participant is seeking for own health issue or other's health issue.

Give the participant an introduction to the eye tracker machine—what it does, how it works, and what the participant should do to obtain accurate results.

Calibrate the eye tracker machine for the participant, following instructions in the eye tracker manual. The calibration is used to adjust the eye tracker system to the participant's gaze positions. The system indicates whether the calibration is successful.

Ask the participant to look for information in the discussion forum that is relevant (or browse for interesting health-related information).

This study did not set a time limit for participants. They kept browsing until they felt satisfied with the information they found or decided to end the session. The average duration of a session is 9.5 minutes (ranging from 4.1 to 14.6 minutes).

Replay a video recording of the participant's search and browse session together with the indications of eye gaze positions and eye movements. Ask the participant to comment on the post surrogates with eye fixation as well as those selected and the text passages in the post content with eye fixation. The purpose was to identify the participant's reasons why a post surrogate or detailed post is relevant or not. More details of the questions are summarized in Table 2.

Table 2. Interview guide to obtain participants' comments on the video recording of his or her eye movement behavior.

Screen	Questions to ask
Summary result screen of post surrogates	
	For post surrogates with eye fixation, ask *What attracted your eyes to these posts? Why is this interesting?*
	For post surrogates clicked on, ask *Why did you select this post for further reading?*
	Look through the post surrogates without eye fixations and point to a few titles that appear similar in topic to the post surrogates the participant had clicked on, and ask *Why was this post not interesting?*
Detailed post screen	
	Point out text segments in the post content with eye fixation and ask *Is this information useful or relevant?*
	Finally, ask *Why do you think this post is relevant or not relevant?*

Study Population

This study was targeted at laymen who did not have severe or critical diseases. A layman refers to normal adults who were not health professionals or did not hold expertise in health-related areas. Their behaviors were thought of as consumer health information behaviors.

They are thought not to have severe or critical diseases because most of the times people who had these diseases would directly refer to the doctors rather than browsing the Internet for related health information. Besides, they are well educated because the text in health forums may require some level of language comprehension that less-educated people are not equipped with.

Sampling Technique

This study took convenient sampling method to recruit research participants. The researchers sent out invitation emails to the students and staffs in Nanyang Technological University and made posters on school canteens and libraries. They also invited their friends by telephone calls and messages and asked them to forward the invitation to their friends. Because it will take about 1 hour to finish the experiment with briefing and post-experiment interview (not including the time for transportation), most of the participants are university students as they have more free time, and it takes just a few minutes' walk in campus for them to get the experiment laboratory.

Sample Size

Overall, 60 study participants were recruited from students and staff of Nanyang Technological University (a large government-funded university in Singapore). An additional 10 participants were recruited from our network of friends who were working adults outside the university. The participants were recruited by advertisements posted on the university campus notice boards and by email, telephone, and oral invitation. The 60 study participants comprised 32 full-time students, 14 part-time students, 4 staff members, and 10 working adults from outside the university.

Two participants failed in the calibration of the eye tracker machine and were excluded from the study, leaving 58 participants who contributed data to the study.

Ethical Consideration

An approval from Institutional Review Board of Nanyang Technological University was obtained before the formal study was carried out as user's eye movements and their feedbacks and comments were used in this study. Besides, they were asked to provide the information about their own problem or their friends and relatives. Hence, an informed consent was asked to sign after the briefing was finished. If they decided not to sign, they would leave this study with no personal information recorded.

Data Analysis and Management

Content analysis of the kind of texts that the participants focused was carried out to find out what kinds of health information they paid attention to and possibly use in the relevance judgment process. In the content analysis, the sentences that participants fixated on were extracted from the screenshots and categorized into different types of health information by 2 coders.

The eye tracker machine recorded participants' eye movements on each screen and generated a static image of each screen with round spots indicating eye fixations and lines indicating quick eye movements, as illustrated in Figure 1. The size of the round spots reflects the relative duration of the eye fixation. Post surrogates and text passages in the detailed posts covered with round spots were considered to be examined and interpreted by the participant and used in the participant's relevance decision. Surrogates and text passages included in the lines of quick eye movements were considered to be skimmed by the participant.

Different types of health information are related to different aspects of health issues, such as symptoms and history of disease related to the user's condition, drug names, and treatments, which can be considered factual knowledge, and diet- and exercise-related information, which are general interest or lifestyle topics. Content analysis of users' commentary provides more support and explanation for the results of the eye movement analysis. We had derived a comprehensive coding scheme in the pilot study to be used in the content analysis of participants' eye movements. The details of derivation are included in Appendix 1 [1,25-27].

Figure 1. Screenshot of detailed post page with participant's eye fixations.
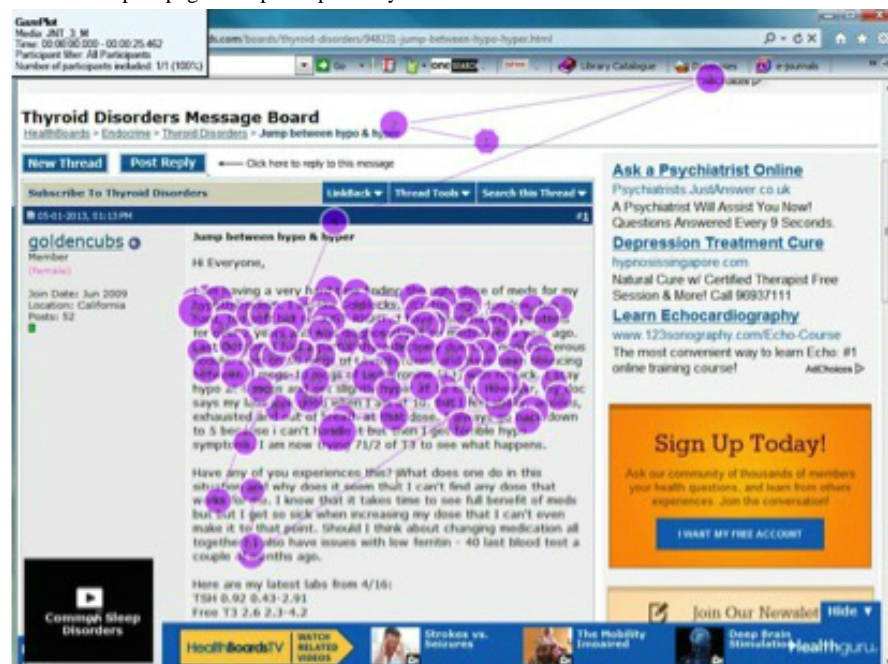
Analysis of Post Surrogates and Post Content With and Without Eye Fixation

The data were downloaded from the eye tracker machine and coded into different types of health information.

Because each participant fixated on and skimmed over different total number of post surrogates and detailed posts, we could not directly use the count of post surrogates and detailed posts containing a particular type of health information for comparison. Instead, we calculated the average percentage of each type of health information found in the post surrogates and the detailed posts for each group of participants. This was computed in 2 steps:

1. Compute the percentage of each type of health information for each participant within a group of participants: count the post surrogates that were fixated on containing a particular type of health information and divide by the overall count of all post surrogates that were fixated on.

2. Compute the average percentage for each group.

The purpose was to find out whether different types of health information were used in the different use contexts and the importance of particular types of health information (ie, relevance criteria) in making relevance judgments.

A screenshot of a detailed post page showing text with eye fixations is shown in Figure 1. The coding was conducted by 2 coders. To check intercoder reliability, both the coders were asked to code about 20% of the data (the data for the first 12 participants). Cohan's kappa between the 2 coders was found to be .82, indicating good consistency. The conflicting codings were discussed with the coders to help them get a clearer conception of the categories. For the purpose of computation, we resolved the conflicting categories by selecting the category

after reviewing with the 2 coders. The remaining 80% of the coding data were divided evenly between the 2 coders.

Analysis of Participant Comments From the Post-Experiment Interview

The participants' comments on whether and why particular post surrogates or post content was relevant were analyzed. The purpose was to elicit the criteria and reasons for their relevance judgments and find out what they thought about the health information on the screen during browsing.

Results

The results of the content analysis indicate that the 3 groups of participants focused on different types of health information: participants seeking for their own health issue focused mostly on symptoms, history of disease, and treatment, which can be considered case-based relevance criteria that participants might use to match their own conditions. Participants seeking for others' health issue focused mostly on treatments, medical terms, and cause of disease, which could be considered general knowledge–based relevance criteria. Finally, participants with no particular health issue in mind focused mostly on general health topics such as diet and exercise, public awareness topics such as smoking and air pollution, and interesting stories. These could be considered general interest–based relevance criteria.

Before the detailed quantitative results of this study, the demographic information of the 58 participants is summarized in Table 3. Most of the participants were Singaporean Chinese (Singapore citizens) and Chinese nationals. The Singaporean participants were all educated in English-medium schools. Chinese nationals were educated in Mandarin in their elementary school in China but took English courses from secondary school to college. Before they came to Singapore, they had to pass 1 or more English tests such as International English Language Testing System and Test of English as a Foreign Language.

Table 3. Summary of demographic information of research participants (N=58).

Characteristic	Demographics	N
Nationality		
	Chinese	31
	Singaporean	24
	Others	4
Degree		
	Undergraduate	15
	Master degree	28
	PhD	15
Occupation		
	Full-time student	30
	Part-time student	14
	Staff	4
	Working adults	10
Age, years		
	18-20	2
	20-30	27
	30-40	25
	40-50	4

Scanning Post Surrogate Page

The percentages of post surrogates containing each type of health information in the scanning stage are shown in Figure 2. Categories with percentages less than 5% were removed from the graph so that it can fit in the page.

It can be seen in Figure 2 that the 3 groups of participants with different types of use contexts focused on different types of health information. For participants seeking for their own health issue, the following were the most common categories they fixated on (the 95% confidence intervals [CIs] are summarized in Table 4):

- Description of patient symptom (B1 SYM: 58.3%).
- Personal history of disease (B3 HST: 57.6%).
- Description of disease (E2 DIS: 55.2%).

In comparison, for the participants seeking for other's health issue, the most common types of information fixated on were as follows:

- Description of disease (E2 DIS: 45.3%)
- Description of terms (E3 TRM: 47.5%)
- Description of procedure used (G2 PRO: 35.8%).

For participants browsing with no particular issue, the most common types of information fixated on were as follows:

- Smoking issue (H3 SMO: 32.4%)
- Hot health topic (H4 HOT: 28.4%; eg, sexual issue, weight control, exercise, diet)

- Rare health issue (I1 RAR: 23.5%).

A few other types of health information were also found important for each group of participants (with frequency above 20%):

- Participants seeking for their own health issue: subjective feeling of having the disease, description of treatment, and disease terminology.
- Participants seeking for other's health issue: type of procedure.

There was a possibility that the percentages for the different types of health information with eye fixation merely reflected the overall distribution in the discussion forum. The *t* test was carried out to compare, for each type of information, the percentage of the surrogates skimmed versus the percentage of the surrogates fixated on. As summarized in Table 4 , the types of health information with higher percentage in the set of surrogates fixated on had significantly lower percentage in the set of surrogates skimmed (eg, SYS symptom 29.3% with a 95% CI of 28.5-29.7 in skimmed surrogates vs 58.3% with a 95% CI of 58.1-58.6 in fixated surrogates). The other types of health information summarized in Table 4 hold the similar patterns. This indicated that the participants with different types of use context intentionally focused on the particular types of health information in making relevance judgments. The other types of health information were not found to be significantly different in fixated and skimmed post surrogates.

Table 4. Percentage of health information in the set of skimmed surrogates compared with the fixated surrogates.

Reason	Category of health information	% In surrogates skimmed (95% CI)	% In surrogates fixated on (95% CI)	*P* value
For their own heath issue (N=18)				
	B1 SYM symptom	29.3 (28.5-29.7)	58.3 (58.1-58.6)	.01
	B3 HST history of disease	18.4 (17.2-18.6)	57.6 (57.2-58.0)	0
	E2 DIS description of disease	29.1 (28.3-29.9)	55 (54.3-55.7)	.02
For other's health issue (N=18)				
	E3 TRM terms	25.1 (24.1-26.1)	47.5 (47.2-47.8)	.01
	E2 DIS description of disease	17.6 (17.3-17.9)	45.3 (45.0-45.6)	.01
	G1 TRT treatment	12.5 (12.0-13.0)	28.4 (28.2-28.6)	.02
No particular issue (N=22)				
	H4 HOT hot topic	16.5 (15.6-17.4)	32.4 (32.0-32.8)	.01
	H3 SMO smoking	17.3 (16.5-18.1)	28.4 (28.0-28.8)	.01
	I1 RAR rare issue	13.2 (12.6-13.8)	23.5 (23.0-24.0)	.01

We further calculated the percentages of different types of health information for the selected post surrogates (ie, clicked on). The selected post surrogates were a subset of the post surrogates with eye fixations. For the selected post surrogates, the most common types of health information found were:

For participants seeking for own health issue

- Description of patient symptom (B1 SYM: 88%)
- Personal history of disease (B3 HST: 75%)
- Description of disease (E2 DIS: 72%).

For those seeking for other's issue

- Descriptions of terms (E3 TRM: 77%)
- Cause of disease (E1 RSN: 72%)
- Description of procedure (G2 PRO: 46.7%).

For those with no particular health issue

- Hot public health topics (H4 HOT: 68%)
- Smoking issue (H3 SMO: 56%)
- Rare health issues (I1 RAR: 28%).

The most common types of information were mostly the same as those for fixated post surrogates but with even higher percentages. The only difference is in the increased importance of cause of disease for participants seeking for others' health issue. The details are shown in Figure 3.

For each type of information that had a high percentage in a particular group (eg, description of symptom for participants seeking for their own issue), we calculated the percentage of fixated post surrogates that were actually selected among these 3 groups. If the participant clicked on the post surrogate, it suggested that the participant thought it likely to be relevant. We calculated the ratio of post surrogates selected over post surrogates fixated on by dividing the number of post surrogates containing the type of information that were selected by the number of post surrogates containing the type of information that were fixated on.

For example, if there were 60 surrogates with fixations containing a description of a symptom and of these 30 were selected, the ratio is 50%. The assumption is that if the participant clicked on a post surrogate, the participant thought the post content likely to be relevant.

As a baseline for comparison, we also calculated the overall percentage of fixated surrogates that were selected by dividing the number of post surrogates that were selected by the number of post surrogates that were fixated on.

The results are summarized in Table 5. Analysis of variance was performed to identify significant differences among the 3 groups of participants, and the *P* values obtained are also summarized in Table 5.

It was found that participants seeking for their own health issue use personal history of disease (*P*=.01), description of patient symptom (*P*=.02), and the description of disease (*P*=.17, but $P_{1,2}<.01$ and $P_{1,3}=.01$) significantly more in determining the relevance of post surrogates. Participants seeking for other's issue used significantly more the treatment information (*P*=.01), cause of disease (*P*=.03), and disease terminology (*P*=.11, but $P_{1,2}$ and $P_{2,3}=.01$) in the relevance judgments. In contrast, participants browsing without a particular issue used a general-interest topic, smoking (*P*=.01), hot topic (*P*=.01), and rare topic (*P*=.02). Thus, participants with different use contexts used significantly different kinds of relevance criteria in their relevance decisions.

Table 5. Percentage of fixated surrogates that were selected for each type of health information.

Type of health information[a]	For own health issue (group 1, N=18), %	For other's health issue (group 2, N=18), %	No particular issue (group 3, N=22), %	ANOVA[b] P-value (sig.)	Post hoc $P_{1,2}$	Post hoc $P_{1,3}$	Post hoc $P_{2,3}$
B1 SYM symptom	37	11	2	.02	.01	.00	.01
B3 HST history	32	14	4	.01	.01	.00	.00
E2 DIS description	30	14	11	.17	.01	.00	.18
E3 TRM terms	7	29	4	.11	.01	.45	.01
G1 TRT treatment	16	35	4	.01	.01	.01	.00
E1 RSN cause	4	45	22	.03	.01	.02	.01
H3 SMO smoking	2	7	52	.01	.01	.00	.00
H4 HOT hot topic	1	6	64	.01	.01	.00	.00
I1 RAR rare	3	10	24	.02	.01	.01	.00
Overall percentage	15	12	8	.37			

[a]Types of health information with low percentage of occurrence in the fixated surrogates are excluded.

[b]ANOVA: analysis of variance.

Figure 2. Percentage of post surrogates containing each type of health information for the 3 groups of participants.
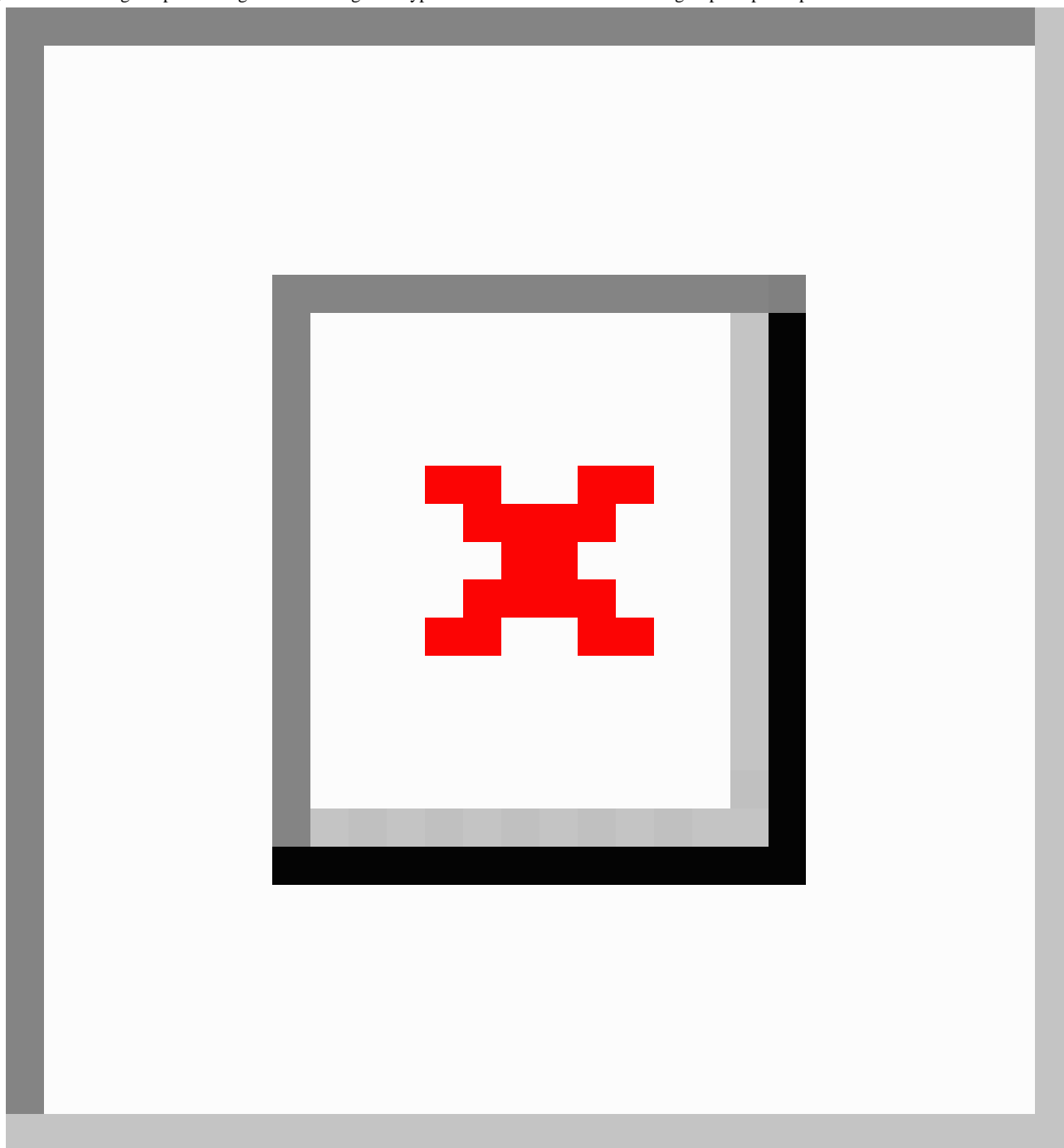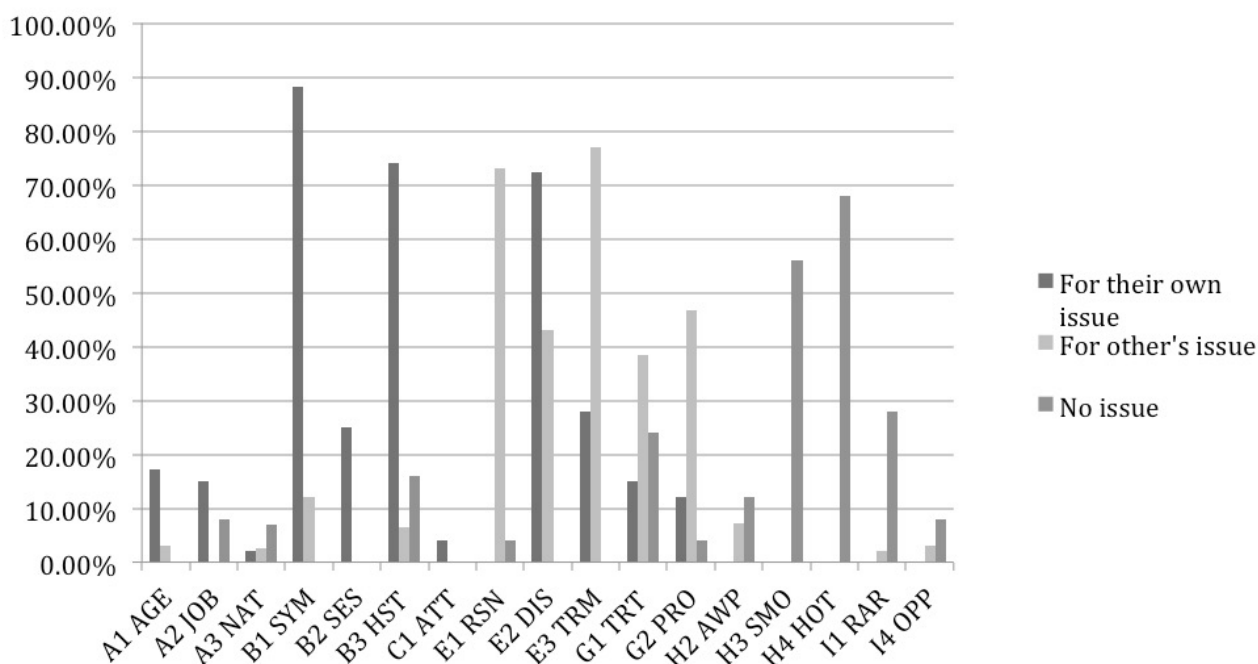
Figure 3. Percentage of selected post surrogates containing each type of health information (the types of health information with less than 5% occurrence are left out).

Reading Post Content Stage

Content analysis was carried out on the posts that were judged relevant by participants, that is, that the participants indicated were relevant during the post-experiment interview. Figure 4 shows the percentages of the different types of information among all the relevant posts.

From Figure 4 , it can be seen that the situation was similar to that of the scanning stage. For people seeking for their own health issue, the most common types of health information that participants fixated on were personal history of disease (91.7%), description of disease (89.6%), and description of patient symptom (87.5%).

For participants seeking for other's health issue, the most common types of health information were description of terms (92.3%), description of procedure (91.5%), and description of treatment (88.7%).

For participants with no particular health issue, the most common types were smoking (66.7%), hot topics (33.7%), and air and water pollution (33.7%).

For comparison, the percentages of posts that were read containing each type of health information are shown in Figure 5. It was found that for participants seeking for others' health issue, the percentages of posts that were read containing personal or case-related types of health information were higher compared with the percentages of posts judged as relevant. In fact, the percentages of posts judged as relevant were nearly

zero, which indicated that participants seeking for others' health issue treated personal details as a sign of nonrelevance. Other than this, the patterns were similar with those for posts judged as relevant.

To further investigate the importance of particular types of health information in determining a detailed post's relevance, we calculated the ratio of posts judged relevant to posts that were read for each type of information by dividing number of posts read and judged relevant, containing this type of information by number of posts read, containing this type of information.

It was found that description of symptoms, personal history of disease, and description of disease were most important types of information for participants seeking for their own health issue (P=.001, .001, and .02) but not for the other 2 groups of participants. For participants seeking for others' health issue, cause of disease, description of terms, and treatment procedure were the important types of information (E1 P=.01; G2 P <.01; E3 P=.23 but $P_{1,2}$=.01 and $P_{2,3}$=.02). In comparison, air pollution, smoking issue, and rare cases were the important types of health information for participants with no issue in mind (P=.02, .01, and .001). The details are summarized in Table 6. These findings were similar to the findings for the scanning stage, suggesting that participants with different types of use contexts used the same types of health information in detailed post reading stage as in surrogate scanning stage as relevance criteria. However, different groups of participants tended to use different types of relevance criteria.

Table 6. Percentages of read posts that were judged as relevant for the most important types of health information.

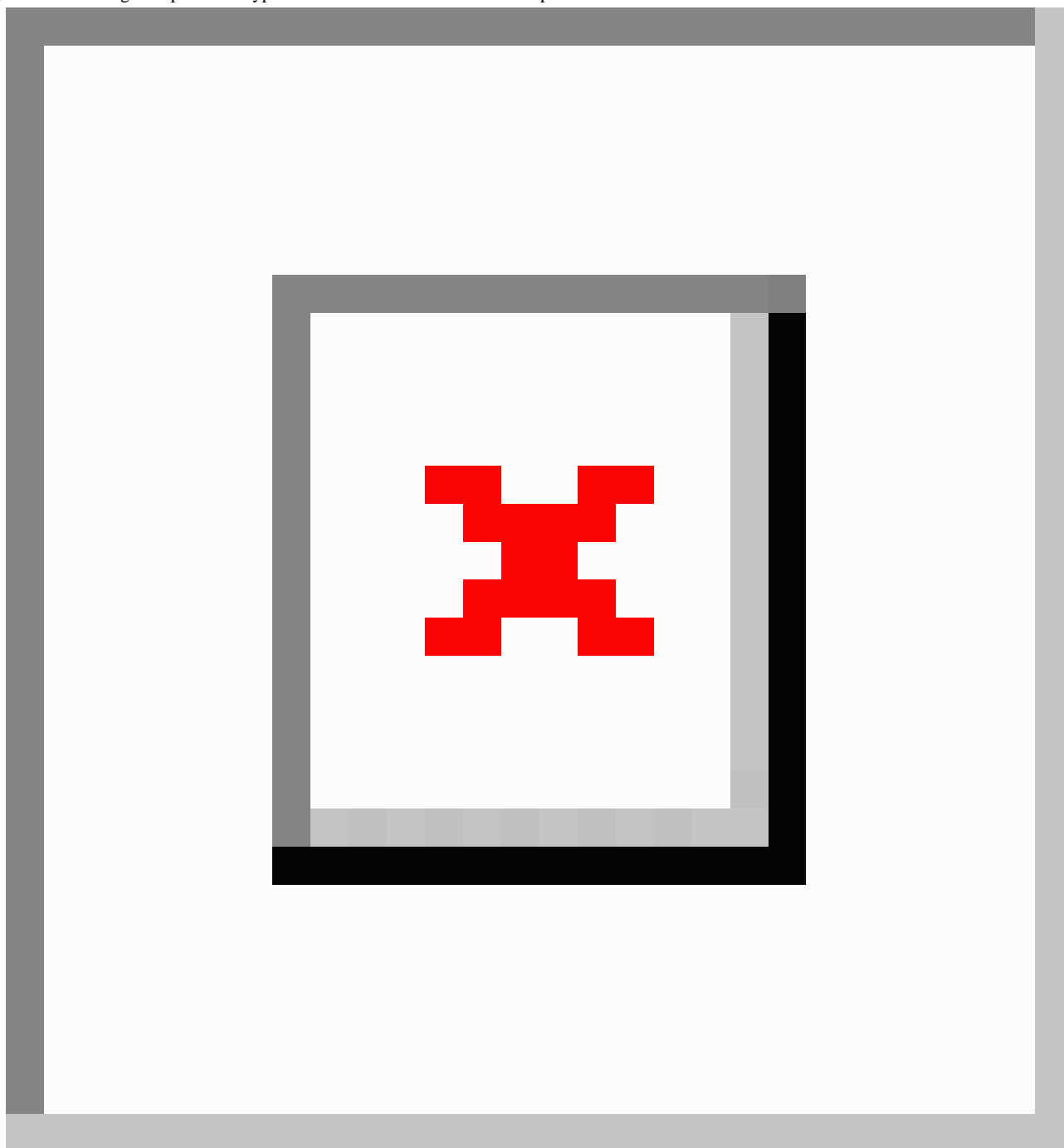| Type of health information	For their own health issue	For other's health issue	With no particular issue	ANOVA[a]P value	Post hoc $P_{1,2}$	Post hoc $P_{1,3}$	Post hoc $P_{2,3}$
B1 SYM symptom	74.5%	32.2%	11.2%	.001	0	.001	0
B3 HST history	72.3%	21.4%	5.6%	.001	0	0	.001
E2 DIS description	69.1%	31.6%	7.1%	.02	.01	.01	0
E1 RSN cause	16.5%	67.2%	9.4%	.01	.01	0	0
E3 TRM terms	6.5%	21.3%	4.5%	.023	.01	.44	.02
G2 PRO procedure	14.4%	53.2%	2.5%	.001	0	0	.001
H2 AWP pollution	4.5%	15.3%	55.9%	.02	.02	0	0
H3 SMO smoking	6.4%	4.5%	45.2%	.21	.67	0	0
I1 RAR rare issue	2.1%	7.5%	42.1%	.001	.001	0	0

[a]ANOVA: analysis of variance.

Figure 4. Percentages of particular types of health information in relevant posts.

Figure 5. Percentages of posts (read) containing each particular type of health information.
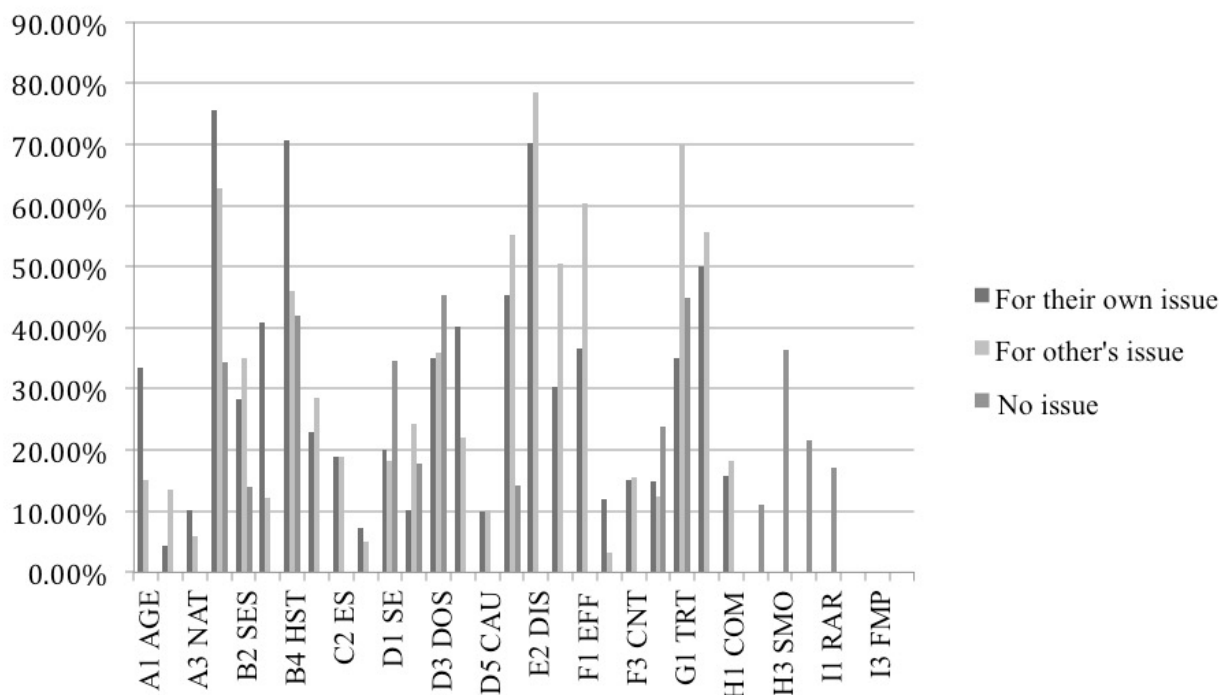
Post-Interview Analysis

The participants were interviewed after the health information seeking session to obtain self-reported commentary on their relevance judgments. Questions were asked to find out what were the reasons that attracted participants' attention and the criteria or reasons why they judged particular post surrogate or detailed post as relevant.

The participants' responses were coded to identify 3 broad types of relevance criteria: case-based criteria, basic or general knowledge–based criteria, and general awareness or curiosity-based criteria. To improve coding consistency, the coding was based on the presence of cue words that reflect the 3 categories. For instance, if a participant's answer included the following terms or phrases "I compare," "similar to my situation," "the similarity between mine and the post," "familiar with the condition/symptom," and other similar phrases, then,

they were coded as case-based criteria. If the answer contained the phrase "I learned before," "I knew it before," "I searched for them in the past," "I heard," or other similar phrases, they were coded as basic or general knowledge–based criteria. The details of the coding scheme are summarized in Table 7.

The coding was done by the first author and another PhD student in the same school (not the same person as the coders for the content analysis of text with eye fixation). The coders looked for cue words or cue phrases as summarized in Table 7 and used them to categorize participants into these 3 groups of criteria. If cue words and cue phrases belonging to more than 1 category were found, then, they were simultaneously counted as belonging to all the matching categories. A high Cohen's kappa of .91 was obtained for the 2 coders. Disagreements between the 2 coders were easily resolved by examining more of the participant's answers.

Table7. Coding scheme for post-experiment interview data.

Table 7.

Criterion	Keywords or phrases to look for
Case-based criteria	
	Compare, match, comparison, similar, consistent
	Similar to my problem, situation, or condition, same as mine or my problem, situation, condition, or disease, like my problem, situation, or condition
Basic or general knowledge–based criteria	
	Know, learn, familiar
	Heard or learn from, read this before from newspaper, book, magazine, or Internet, somebody told or informed me
General awareness or curiosity-based criteria	
	Interesting, funny, rare, weird, strange, curious
	Smoking, pollution, exercise, diet, fitness
	Everybody know, hot topic, public concern, environmental, healthy, well-being

Some participants provided detailed explanations of why they judged 1 particular post surrogate or detailed post as relevant. They mentioned in which ways they compared the post with their own conditions, such as age, location of problem, description of the pain, and diagnosis result. However, most of the participants provided only basic reasons. On the basis of the participant's initial answer, we followed up with more probing questions. As the interview was conducted after the experiment session, the participants were somewhat tired and not many gave very detailed explanations.

From the analysis of the answers, it was found that for participants seeking for their own health issue, 13 of 18 (72%) clearly used condition or symptom match in their judgment of post surrogates and detailed posts. Here is a quote from a participant:

> I found the guy in this post was quite similar to me in the position of headache. He also had the problem after a long time of sitting in front of the screen, the same as me. He did some massage but didn't work for a long time. I also tried this but found not very useful.

Three participants were ambiguous in their explanation but mentioned some comparison between themselves and others:

> When I read the post I realized the boy did not have a history of jogging as me. His pain was due to the injury of his leg. I do not think it is the same reason as me since my pain comes from the muscle.

The remaining 2 participants expressed other reasons for judging relevance. One wanted to find new information he did not know before. Another used the reference from well-known medical journals or websites as the criterion of relevance.

For participants seeking for other's health issue, 15 of 18 (83%) used medical and health knowledge that they gained from school or previous searching to guide them in the relevance judgment. For example:

> I heard from my friend about acid reflux. So I want to know more about this and can share with him what I find here.

Three of them made the judgment with other reasons. Two mentioned novelty of the information as the reason they judged the post as relevant. The last participant did not trust information from general users of this health discussion forum and wanted to find comments from health professionals.

For participants with no particular issue, 16 of 22 (73%) participants was found to judge relevance of post surrogates and detailed posts by their interest, curiosity, rarity, and some well-known health issues (eg, diet and exercise). For example:

> I found it funny since I did not expect that a girl looking to lose some weight would believe that only magic pills can help her. She did not do any exercise and that is impossible.

The other cluster of 6 of 22 (27%) participants judged relevance by a mixture of case-based matching and basic knowledge criteria. Moreover, it was found that these participants actually had latent health information needs. For some posts, they used case-based criteria similar to the participants seeking for their own health issue. For other posts, they used their prior medical knowledge. We labeled these as "participants with latent health information needs."

Latent health information needs were detected when these participants were asked the reason for reading particular detailed posts. Because they were grouped as participants with no particular health issue, they were expected to respond with reasons related to personal interest and curiosity. Instead, their answers suggested that they had some health issues related to the topics of these detailed posts or had heard from their friends about a health issue. For example:

> When I come across the post, I recall I sprained my left foot three years ago. When I read the post, I found that the guy sprained his foot because of running, the same as me. So, I continued to read and found it relevant. I can make use of his experience to avoid sprain in the future.

> Actually my grandma had diabetes, so I want to know if any useful information in the post. I realized the man in the post used diet control and I heard before.

> *Some reply suggests some herb medicine will help the patient to control the level of blood glucose. I think it useful and will tell my grandma.*

Their responses indicated that latent health information needs were activated at some point during the information-seeking session when they came across familiar topics that triggered a memory.

Another possibility for the observed latent health information needs is that the participants were reluctant to share their real health information needs with the researcher at the beginning of the experiment. They became more comfortable later after reading posts of the health issue. However, we did not detect clear evidence of this in the study.

Discussion

It was found in this study that users browsing a health discussion forum made use of different types of information in relevance judgments and exhibited different eye movement patterns in the 3 types of use contexts—seeking information for self, seeking for others, and browsing with no particular issue in mind. Users browsing for their own health issue were found to use mainly case-based relevance criteria such as symptoms, personal history of disease, and description of disease and personalized treatment in their judgments. Participants seeking for others' health issue were found to use mainly general knowledge–based criteria such as medical terms, cause of disease, and basic treatments and procedures in their judgments. In contrast, participants seeking with no particular health issue were found to be interested in general health topics, hot topics, and rare health issues.

The personalized treatment refers to the customized treatments that patients received based on their unique conditions, whereas basic treatment refers to the general treatment that can be found in medical books or manuals.

Looking at the results in more detail, participants seeking for own health issue focused mainly on the poster's symptoms, personal history of disease, and description of disease both when scanning post surrogates and reading detailed post content. These case-based criteria could be considered as more detailed categories of the comparison relevance category identified by Huang and Soergel [12]. They defined comparison relevance as the relevance derived from the similarity between 2 different cases. People who are seeking for their own health issue often compare their own situation (experiences, feelings, symptoms, and history of disease) with the description of the poster's situation.

Participants seeking for other people's health issue focused on the terminology, description of the disease, cause of disease, and available treatments. These types of information can be considered to be more detailed relevance criteria within the broader category of content or information in the framework of Cool, Belkin, and Kantor. The category of content or information can be interpreted as the factual knowledge of health issues and treatments. When people seek health information for others' health issue, they are often not familiar with the details

of the patient and have to consider only generic medical information in their relevance judgment.

It cannot be concluded that the relevance criteria identified in this study are the only ones used when browsing a health discussion forum. People also use other relevance criteria implicitly, such as topicality, accuracy, presentation, and authority, but, these cannot be determined just from content analysis of text with eye fixation. People seeking for their own issue and other's health issue must find the right topic before they can check and read other details. Participants sometimes check the poster's profile, which suggests that authority is also considered. It is likely that these implicit relevance criteria will become the focus of attention when they are violated. However, no instance of this was encountered in this study. These criteria might also be important considerations when users decide to actually use the information or adopt a recommendation in practice.

The results of this study carry implications for the design of more user-oriented health information systems. The relevance ranking by the search function can assign different weights to different types of information, depending on whether the user is searching for self, others, or with no particular issue in mind. The user can be prompted to select one of these use contexts when accessing the system. For users seeking for their own health issue, the summary result page can display the post surrogates that best match the user's profile, if available. If the user is browsing with no particular topic in mind, the summary result page can display post surrogates that were clicked on and viewed by the highest number of previous users, indicating topics of general interest. A metadata field for posters to indicate the type of health information included in their post can help the system to filter and display posts that better match the use context and health profile of the user.

Limitations

This study has some limitations that need further exploration and investigation:

The participants are residents of Singapore (most of them are Singaporeans and Chinese nationals). This study did not include participants from other ethnic groups and nationalities.

The participants did not have a critical health problem at the time of conducting the study. People with severe problems may exhibit different relevance judgment behaviors.

The demographic profile of participants did not cover all segments of the society. In particular, the participants were either undergraduate students or had at least an undergraduate degree.

This study did not consider the influence of human factors, such as personality and attitude to the Internet. Two participants were found to have long examining duration owing to their reading habit developed in childhood.

The health information used in this study was written in English. Content in other languages might influence people's eye movements and corresponding relevance judgments. For example, Chinese characters are quite different from English text in size and ways of organization.

Conclusions

This study chose a particular health discussion forum—HealthBoards [6]—as the study platform. It is representative of user-contributed content in health discussion forums, as they are similar in structure and content. Hence, the results of this study are very likely to hold with other Web-based health discussion forums. In addition, the results might also be applicable to other types of social media websites (eg, Facebook groups for various diseases) with lots of user-contributed content. People seeking for their own health issue might look for Facebook or blog pages of people with the most similar condition (ie, identify the most similar person rather than the most similar post), as information on Facebook and in blogs is organized by person rather than topic.

In contrast, the results may not be applicable to authoritative health websites maintained by health care organizations and government agencies (eg, Mayo Clinic and PubMed). People seeking for their own health issue on these websites may have difficulty matching their own condition with the description on the websites. They have to use basic medical terms in searching and making relevance judgments.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The derivation of coding scheme.

[PDF File (Adobe PDF File), 47KB - jmir_v18i6e136_app1.pdf]

Multimedia Appendix 2

Health discussion forum candidates searched through Google search engine.

[PDF File (Adobe PDF File), 13KB - jmir_v18i6e136_app2.pdf]

References

1. Pew Research Center. 2013. Health online URL: http://www.pewinternet.org/2013/01/15/health-online-2013/ [accessed 2015-12-27] [WebCite Cache ID 6e5d1gREg]
2. Barry C. User-defined relevance criteria: an exploratory study. J Assco Inf Sci Technol 1994;45(3):149-159.
3. Schamber L. Users' criteria for evaluation in a multimedia environment. 1991 Presented at: Proceedings of the 54th ASIS Annual Meeting; Oct; 1991; Washington, DC p. 37-31.
4. Balatsoukas P, Ruthven I. An eye-tracking approach to the analysis of relevance judgments on the Web: The case of Google search engine. J Am Soc Inf Sci Tec 2012 Aug 17;63(9):1728-1746. [doi: 10.1002/asi.22707]
5. Kim S, Oh S. Users' relevance criteria for evaluating answers in a social Q&A site. J. Am. Soc. Inf. Sci 2009 Apr;60(4):716-727. [doi: 10.1002/asi.21026]
6. HealthBoards. URL: http://www.healthboards.com/boards/index.php [accessed 2016-01-02] [WebCite Cache ID 6eEsVFcOx]
7. Lambert SD, Loiselle CG. Health information seeking behavior. Qual Health Res 2007 Oct;17(8):1006-1019. [doi: 10.1177/1049732307305199] [Medline: 17928475]
8. Lenz ER. Information seeking: a component of client decisions and health behavior. ANS Adv Nurs Sci 1984 Apr;6(3):59-72. [Medline: 6426379]
9. Carron-Arthur B, Ali K, Cunningham JA, Griffiths KM. From Help-Seekers to Influential Users: A Systematic Review of Participation Styles in Online Health Communities. J Med Internet Res 2015;17(12):e271 [FREE Full text] [doi: 10.2196/jmir.4705] [Medline: 26627369]
10. Taylor AR, Cool C, Belkin NJ, Amadio WJ. Relationships between categories of relevance criteria and stage in task completion. Information Processing & Management 2007 Jul;43(4):1071-1084. [doi: 10.1016/j.ipm.2006.09.008]
11. Saracevic T. Relevance: A review of the literature and a framework for thinking on the notion in information science. Part III: Behavior and effects of relevance. J. Am. Soc. Inf. Sci 2007 Nov;58(13):2126-2144. [doi: 10.1002/asi.20681]
12. Huang X, Soergel D. Inform Research(1)?. 2007. An evidence perspective on topical relevance types and its implications for exploratory and task-based retrieval URL: http://www.informationr.net/ir/12-1/paper281.html [accessed 2015-12-27] [WebCite Cache ID 6e5jaunEZ]
13. Miwa M, Egusa Y, Saito H, Takaku M, Terai H, Kando N. Information Research(3). 2011. A method to capture information encountering embedded in exploratory Web searches URL: http://www.informationr.net/ir/16-3/paper487.html [accessed 2015-12-27] [WebCite Cache ID 6e5k3zRzZ]
14. Rayner K. Eye movements in reading and information processing: 20 years of research. Psychol Bull 1998 Nov;124(3):372-422. [Medline: 9849112]

XSL·FO
RenderX

15. Bidmon S, Terlutter R. Gender Differences in Searching for Health Information on the Internet and the Virtual Patient-Physician Relationship in Germany: Exploratory Results on How Men and Women Differ and Why. J Med Internet Res 2015;17(6):e156 [FREE Full text] [doi: 10.2196/jmir.4127] [Medline: 26099325]

16. Jamal A, Khan S, AlHumud A, Al-Duhyyim A, Alrashed M, Bin SF, et al. Association of Online Health Information-Seeking Behavior and Self-Care Activities Among Type 2 Diabetic Patients in Saudi Arabia. J Med Internet Res 2015;17(8):e196 [FREE Full text] [doi: 10.2196/jmir.4312] [Medline: 26268425]

17. Song H, Omori K, Kim J, Tenzek K, Hawkins J, Lin W, et al. Trusting Social Media as a Source of Health Information: Online Surveys Comparing the United States, Korea, and Hong Kong. J Med Internet Res 2016;18(3):e25 [FREE Full text] [doi: 10.2196/jmir.4193] [Medline: 26976273]

18. Puspitasari I, Moriyama K, Fukui K, Numao M. Effects of individual health topic familiarity on activity patterns during health information searches. JMIR Med Inform 2015;3(1):e16 [FREE Full text] [doi: 10.2196/medinform.3803] [Medline: 25783222]

19. Tang R, Solomon P. Use of relevance criteria across stages of document evaluation: On the complementarity of experimental and naturalistic studies. J Am Soc Inf Sci 2001;52(8):676-685. [doi: 10.1002/asi.1116]

20. Crystal A, Greenberg J. Relevance criteria identified by health information users during Web searches. J Am Soc Inf Sci 2006 Aug;57(10):1368-1382. [doi: 10.1002/asi.20436]

21. Buscher G, Dengel A, van Elst L. Eye movements as implicit relevance feedback. In: CHI EA '08 CHI '08 Extended Abstracts on Human Factors in Computing Systems. New York, NY: ACM; 2008:2991-2996.

22. Zhang Y. Consumer health information searching process in real life settings. Proc Am Soc Info Sci Tech 2013 Jan 24;49(1):1-10. [doi: 10.1002/meet.14504901046]

23. Saracevic T. Relevance reconsidered. 1996 Presented at: Proceedings of the Second Conference on Conceptions of Library and Information Science; 1996; Copenhagen, Denmark.

24. Spink A, Greisdorf H, Bateman J. From highly relevant to not relevant: examining different regions of relevance. Information Processing & Management 1998 Sep;34(5):599-621. [doi: 10.1016/S0306-4573(98)00025-9]

25. Kishimoto K, Fukushima N. Use of anonymous Web communities and websites by medical consumers in Japan to research drug information. Yakugaku Zasshi 2011;131(5):685-695 [FREE Full text] [Medline: 21532265]

26. Cotten SR, Gupta SS. Characteristics of online and offline health information seekers and factors that discriminate between them. Soc Sci Med 2004 Nov;59(9):1795-1806. [doi: 10.1016/j.socscimed.2004.02.020] [Medline: 15312915]

27. Chuang KY, Yang CC. Interaction patterns of nurturant support exchanged in online health social networking. J Med Internet Res 2012;14(3):e54 [FREE Full text] [doi: 10.2196/jmir.1824] [Medline: 22555303]

Abbreviations

A1 AGE: age and gender
A2 JOB: job and occupation
A3 NAT: nationality
B1 SYM: description of patient symptom
B2 SES: subjective feeling of having a problem
B3 HST: personal history of disease
C1 ATT: attitude to the problem
C2 ES: emotional status of knowing the problem
C3 OA: other's attitude and support
D1 SE: perceived side effect
D2 INT: interaction with another health problem
D3 DOS: dosage used
D4 USE: description of procedure used
D5 CAU: caution and reminder
E1 RSN: cause of disease
E2 DIS: description of disease
E3 TRM: description of terms
F1 EFF: efficacy
F2 IND: indications
F3 CNT: contraindications
F4 INT: interaction with other drugs
G1 TRT: description of treatment
G2 PRO: description of procedure
H1 COM: common health issue
H2 AWP: air and water pollution
H3 SMO: smoking

H4 HOT: hot topics
I1 RAR: rare health issue
I2 STO: interesting story
I3 FMP: famous person
I4 OPP: counter-intuitive information

Original Paper

Designing Health Websites Based on Users' Web-Based Information-Seeking Behaviors: A Mixed-Method Observational Study

Patrick Cheong-Iao Pang[1,2], MSc; Shanton Chang[1,3], PhD; Karin Verspoor[1,3], PhD; Jon Pearce[1], PhD

[1]Department of Computing and Information Systems, The University of Melbourne, Parkville, Australia

[2]NICTA Victoria Research Lab, Melbourne, Australia

[3]Health and Biomedical Informatics Centre, The University of Melbourne, Melbourne, Australia

Corresponding Author:
Shanton Chang, PhD
Health and Biomedical Informatics Centre
The University of Melbourne
Level 1, 202 Berkeley St
Melbourne, 3010
Australia
Phone: 61 3 83441583
Fax: 61 3 93494596
Email: shanton.chang@unimelb.edu.au

Abstract

Background: Laypeople increasingly use the Internet as a source of health information, but finding and discovering the right information remains problematic. These issues are partially due to the mismatch between the design of consumer health websites and the needs of health information seekers, particularly the lack of support for "exploring" health information.

Objective: The aim of this research was to create a design for consumer health websites by supporting different health information–seeking behaviors. We created a website called *Better Health Explorer* with the new design. Through the evaluation of this new design, we derive design implications for future implementations.

Methods: Better Health Explorer was designed using a user-centered approach. The design was implemented and assessed through a laboratory-based observational study. Participants tried to use Better Health Explorer and another live health website. Both websites contained the same content. A mixed-method approach was adopted to analyze multiple types of data collected in the experiment, including screen recordings, activity logs, Web browsing histories, and audiotaped interviews.

Results: Overall, 31 participants took part in the observational study. Our new design showed a positive result for improving the experience of health information seeking, by providing a wide range of information and an engaging environment. The results showed better knowledge acquisition, a higher number of page reads, and more query reformulations in both focused and exploratory search tasks. In addition, participants spent more time to discover health information with our design in exploratory search tasks, indicating higher engagement with the website. Finally, we identify 4 design considerations for designing consumer health websites and health information–seeking apps: (1) providing a dynamic information scope; (2) supporting serendipity; (3) considering trust implications; and (4) enhancing interactivity.

Conclusions: Better Health Explorer provides strong support for the heterogeneous and shifting behaviors of health information seekers and eases the health information–seeking process. Our findings show the importance of understanding different health information–seeking behaviors and highlight the implications for designers of consumer health websites and health information–seeking apps.

XSL•FO
RenderX

Introduction

The Internet has been widely used for accessing health information [1]. However, finding and discovering the right and useful information remains problematic [2-5]. Despite the efforts made to improve the provision of online health information, some complications still exist for the health information seeker, such as, inefficiently using search engines [6-9], lacking the cognitive skills and health literacy for searching [10-13], and feeling disengaged with health websites due to usability and design issues [14,15]. These problems add burdens to health information seekers and reduce the overall user experience.

As discussed in prior work, one of the reasons behind these problems is that the design of health websites does not address the needs of health information seekers [16-19]. Depending on the specific scenario, seekers may demonstrate either focused or exploratory search approach in the information seeking process [17,20,21]. With different search approaches, seekers use different strategies to find information, which need to be supported with features such as tools for exploratory search, reading-friendly user interface, and memory aids for reviewed information [19]. These features or considerations are often missing in the design of health websites and health information–seeking apps.

To understand and support health information–seeking behaviors, a website called Better Health Explorer (BHX) was designed using a user-centered design approach. The design is based on the conceptualization of search approaches and a classification of health information–seeking behaviors [19]. In this paper, we aim to investigate what improvements can be made to the current tools for health information seeking, by using BHX as the vehicle of a human-based evaluation.

This project adopted a mixed-method approach. Overall, 31 participants took part in an observational study of using BHX and an existing live health website. Overall, our proposed design shows a positive result in improving the user experience of health information seeking. The results show better knowledge acquisition, a higher number of page reads, and more query reformulations in different search tasks. Moreover, participants spend more time to discover health information in exploratory search tasks with our design. In addition, we summarize 4 design considerations for designing consumer health websites and health information–seeking apps, namely: (1) providing a dynamic information scope; (2) supporting serendipity; (3) considering trust implications; and (4) enhancing interactivity.

We demonstrate that BHX improves the support for heterogeneous and shifting behaviors of health information seekers and eases the health information–seeking process. We therefore recommend that designers, Human-computer interaction (HCI) practitioners, and researchers consider these findings in the design of consumer health websites and health information–seeking apps.

Methods

This section will discuss the different aspects of our study. First, we will introduce the rationale and the implementation of BHX, which was the website used in our study. Then, we will explain the design of our user study and finally, the methodology for data analysis.

Better Health Explorer

Better Health Explorer is not only a new user interface (UI) for health information seeking but also a vehicle for evaluating the design and highlighting the considerations for designing consumer health websites. In this subsection, we will start with a brief introduction of the theoretical framework that underpins the design, followed by a short description of UI features that reflect this framework in the implementation of BHX.

Theoretical Background

The design process of BHX follows a user-centered approach. Research has proposed to use user-centered design for eHealth technologies and informational websites [22,23]. In our prior work [17-19], we have conceptualized both the search approaches used in health information seeking and the health information–seeking behaviors demonstrated by different seekers. We have learnt that seekers adopt differing search approaches for different scenarios, and thus, the selected search approach affects the actual health information–seeking behavior. This theoretical work underpins the design of BHX. The following paragraphs will give a brief explanation about this conceptualization.

Focused and Exploratory Search

We have identified 2 search approaches in the process of health information seeking, namely focused search and exploratory search [24,25]. In the health context, we have found that people demonstrate both types of search approaches depending on a number of factors, for instance, their level of knowledge about the health problem, their levels of curiosity, the perceived situational relevance to the health problem, and so forth [17]. Figure 1 illustrates the 2 search approaches and their differences.

Searching with precise keywords and iteratively narrowing down the search scope are sample activities of focused search. As shown in Figure 1, focused seekers concentrate on a small range of information. They often have better knowledge about the health problem and a clearer idea of what they are looking for [17]. However, this approach is difficult for many seekers, as not everyone can clearly express the health issue through search queries or use accurate terminology [7,8].

Exploratory search is another search approach found in health information seeking. Exploratory seekers often are unfamiliar with the knowledge domain and feel unsure about the search goal [24,26]. In these cases, exploratory search often arises, along with an unclear search target, as well as a wider and sparse search scope [17,19], as illustrated in Figure 1. Exploratory search also introduces learning and investigative activities beyond simply finding particular information [25], to clarify the problem and gain an overview of the situation [17].

Figure 1. An illustration of focused and exploratory search approaches.
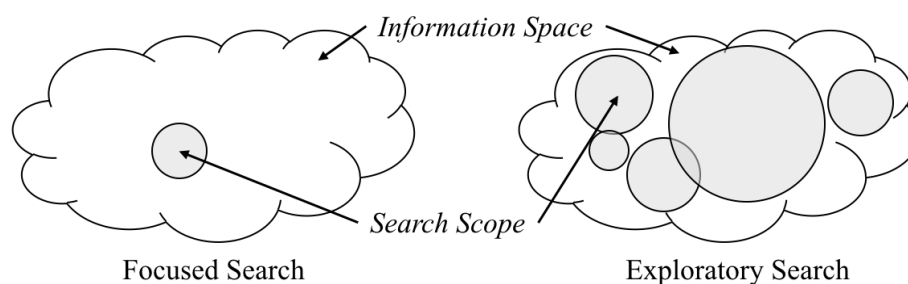
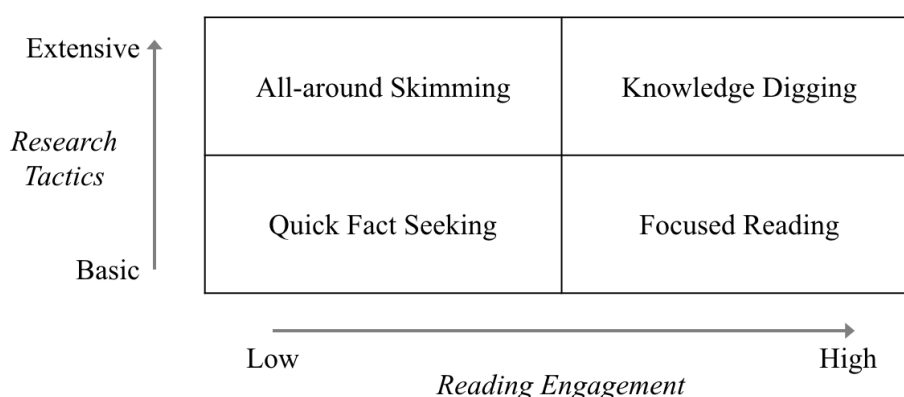### Health Information–Seeking Behavior

Motivated by diverse scenarios and different search approaches, health information seekers implicitly expose 4 different behaviors. As seen in Figure 2, health information–seeking behaviors can be represented with 2 dimensions: Reading Engagement and Research Tactics [19]. Reading Engagement states that a seeker prefers to commit either long or short time for reading, whereas Research Tactics captures that a seeker intends to gain a comprehensive understanding or merely seeks basic facts about the health problem.

This classification is not used for dividing individual health information seekers into different groups but for understanding their behaviors as a whole. Therefore, Figure 2 shows the range of potential behaviors users might engage in when seeking online health information. This classification can assist designers in making apps that are sensitive to the variety of user behaviors as it provides a lens for understanding all the possible behaviors that can be observed within a health website. In addition, this information is important for creating UI elements for different types of searches and providing suitable information for the diverse behaviors. Nevertheless, the motivations and the context of individual behavior are not the focus of this research.

From the design perspective, each of these behaviors leads to different requirements for user interactions with health websites. For example:

- *Quick Fact Seeking* refers to retrieving the superficial information for a specific health topic and terminating the search once it is found. For this type of behavior, websites should provide key points and a brief summary relevant to the topic.
- *All-Around Skimming* goes through a wide range of information in a fast manner. Excerpts and previews will be helpful to support this behavior, for determining what content is useful within a potentially large number of search results.
- *Focused Reading* denotes concentrated reading on a particular topic. As lengthy reading is involved in this case, reader-friendly features (eg, larger font size, bookmarking, highlighting, and so forth.) are recommended to support this behavior.
- *Knowledge Digging* indicates the intense reading associated with the in-depth research on a number of diverse health topics. Providing a broader range of information can assist users to investigate from multiple perspectives.

With this model, UIs can be designed and built by understanding the user actions associated with these behaviors [18,27].

Figure 2. The classification of health information–seeking behaviors.

Changeable Search Approaches and Behaviors

Health information seekers do not adhere to a single search approach or behavior in their search processes. Instead, they choose the approach that is most appropriate to them, based on the circumstances, the context, the urgency of the health issue,

the situational relevance, and personal preferences [17,19,28]. Therefore, it is not feasible to design for just a single or a limited number of seeker types, but rather, we should focus on the properties of each health information–seeking behavior, and support such properties through the design. In this way, the

design will cover most of the actions executed by each type of health information–seeking behavior.

In this subsection, we have introduced different health information–seeking behaviors and their implications for website design. The next subsection will present the actual website implementation, guided by this theoretical framework.

Design Considerations

Rather than replicating the work of sophisticated keyword-based search engines, the design goals of BHX are to address the needs of health information seekers and to deliver an interactive and engaging experience in the health information–seeking process. We will introduce briefly about this website in the following section. Readers can obtain the details of BHX features in another paper [18]. A video demonstration of BHX can also be found in (Multimedia Appendix 1).

Figure 3 displays a screenshot of BHX. The UI looks similar to an ordinary health website, but the exploration panel on the right gives a different experience of finding and exploring health information. Information exploration is facilitated by the list of *tiles* (top right of the screen) and *sliders* (bottom right of the screen).

Sliders are used for generating and refining queries. Health information articles that match the criteria are displayed as colored tiles at the top right of the screen. Colors denote the category that the information belongs to. Therefore, the color pattern offers an overview of the composition of the results. The keyword-less approach brought by the sliders can reduce the cognitive load of looking for health information.

In prior studies, we have learnt that health information seekers use mainly 4 different criteria to seek information [17,19]. Four sliders corresponding to these criteria are provided. In this way, the website can provide a broad range of information with hundreds of combinations of slider values based on the context currently being viewed. This satisfies the needs of different search approaches and behaviors. Meanwhile, the elimination of keyword search resolves the difficulties of generating new search queries.

Common UI elements can assist in health information seeking. The summary and the table of contents can provide an overview and structure of an article to seekers. These features are useful for behaviors of low Reading Engagement. Besides, the "breadcrumb" history bar can reflect the initial goal and the path of the search session. This allows backtracking quickly, understanding, and adjusting the goal in the progress of the information exploration [24].

An example of a concerned mother can further explain the innovative experience of seeking health information. The mother wants to find some information about a common cold for her child, and she looks for it in BHX. After some initial reading about colds, she desires to learn more around this topic and therefore adjusts the sliders. As the sliders change, the tiles start to move and jostle their positions. Her attention is grabbed by an item labeled "pneumonia". She clicks on it, and the reading area is updated with information about pneumonia. Meanwhile, a new set of tiles pops up based on the new topic (pneumonia), so that she can explore further. The "journey" of exploration repeats again with similar steps.

Figure 3. The user interface of Better Health Explorer.
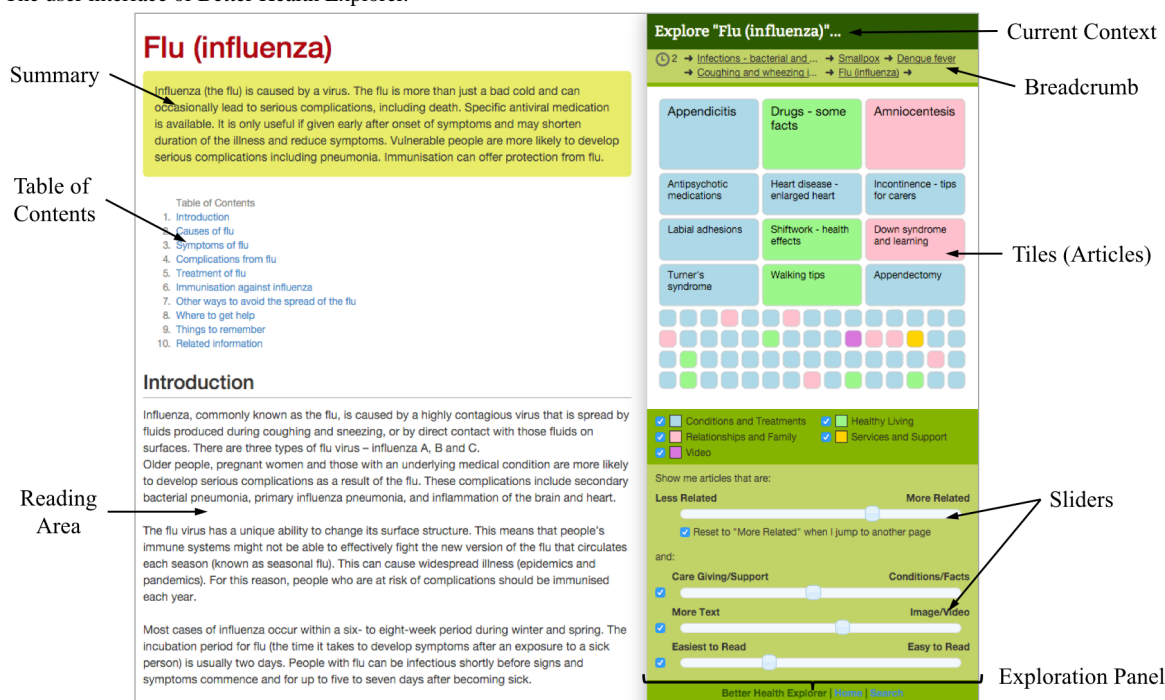
Content

The content used in BHX is obtained from Better Health Channel (BHC) [29], which is a consumer health website established by the Victorian State Government in Australia.

The dataset covers more than 250 health and medical topics for the general public. In addition to text-based materials, pictures and figures from the BHC site, as well as video clips published in their YouTube channel, are included in this study.

Study Design

The study aims to understand how the design of BHX supports health information seeking and information exploration within consumer health websites and to summarize design considerations for future websites. We used a mixed method approach for this study, capturing both qualitative and quantitative data. This research was approved by the university's human ethics committee.

Volunteers were invited to use BHX in a laboratory-based observational study from September to October 2015. For the sake of comparison, they also used BHC, which is an existing live website and also the source of health-related content in this study. Figure 4 shows the UI of BHC. Data were recorded from the participants' use of each website, BHX and BHC, for the assigned study tasks, to assess the relative impact of the BHX design. The design of this study is similar to that of a number of other studies on health information seeking, which conducted observational studies [11,15,30-32].

Figure 4. The appearance of the baseline website used in this study.
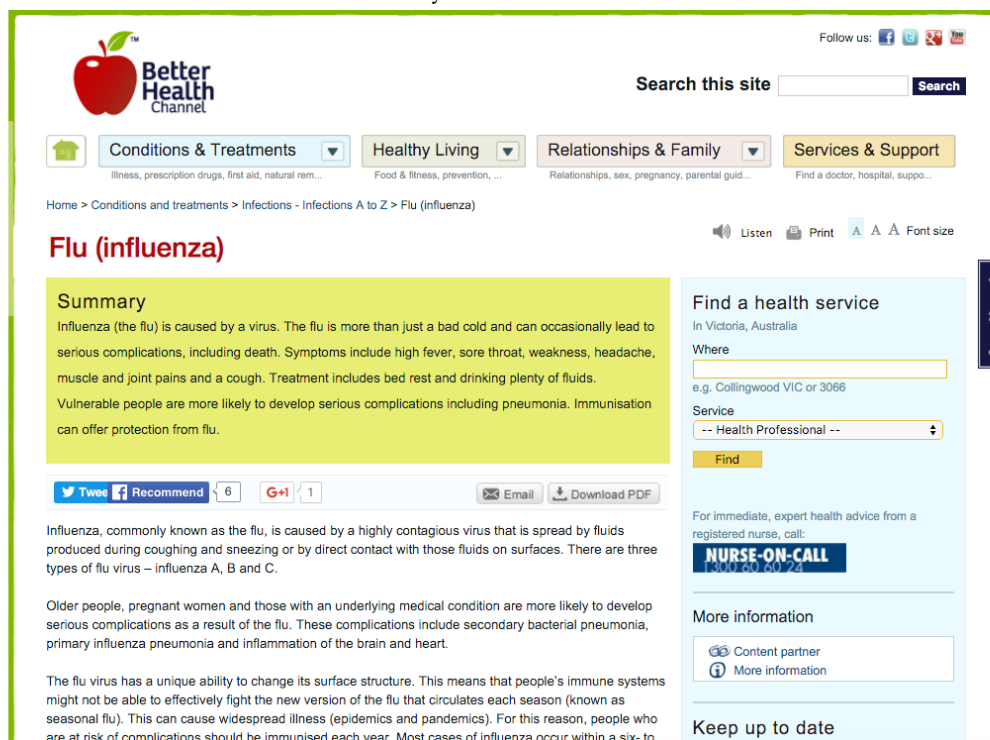
Search Tasks

Participants were given 4 tasks consisting of 2 focused and 2 exploratory search tasks (Table 1), to observe the differences between the 2 search approaches. The tasks posed fictional health scenarios and questions about the scenarios, and we asked the participants to find answers for these questions from either website. We varied the health conditions in the task descriptions that were assigned to individual participants for their search tasks (variables underlined in Task A and B), avoiding conditions that participants may have prior knowledge of, based on their self-reported information (Multimedia Appendix 2). This particular setup was to avoid repeats when testing different websites and to minimize the potential impact of prior knowledge affecting the outcomes, for example, if the participants had suffered from that particular sickness before.

Table 1. Task descriptions used in the study.

Task	Website used	Type[a]	Description
A	Better Health Channel (Baseline)	F	Imagine one of your family members has recently been diagnosed with Type 2 Diabetes/hypertension. As you're living together, your daily life might need to be changed in different ways as well. Please identify three kinds of changes that might be needed in your everyday life.
B	Better Health Explorer	F	Imagine one of your close friends has recently been diagnosed osteoporosis/asthma. As you're living together, your daily life might need to be changed in different ways as well. Please identify three kinds of changes that might be needed in your everyday life.
C	Better Health Channel (Baseline)	E	Imagine you are going to a party and will discuss health information with your friends. Use the website provided by us to identify some interesting health topics. Continue reading until you think it is enough for the discussion.
D	Better Health Explorer	E	Imagine you are going to a party and will discuss health information with your friends. Use the website provided by us to identify some interesting health topics. Continue reading until you think it is enough for the discussion. (Same as Task C)

[a]E: exploratory search task; F: focused search task.

The design of these tasks was purposefully considered. The health issues used in the task descriptions were sourced from the most popular searched keywords from the BHC website. This ensured that the health issues were common and realistic, and the data source contained a substantial amount of information for searching. Also, the contrastive setup with 2 search approaches could facilitate the observation of their differences [33].

The design of exploratory tasks followed the principles outlined in a study by Wildemuth and Freund [34]. The scenario involving social discussions was also found to be helpful for generating exploratory search [35]. In addition, the motivation of searching for family members (Task A) and close friends (Task B) is similar as the perceived situational relevance to the seeker is high in both the cases [17].

Procedure

Participants began with a brief introduction to the study and the 2 websites used, followed by a demographic questionnaire. Informed consent was obtained through a signed consent form. At this stage, participants had to select one of the tasks that they had no prior experience with. Search tasks were then conducted on a desktop computer.

Participants sequentially carried out the search tasks on a desktop computer in a defined order. For counterbalancing learning and ordering effects, the order of the tasks was allocated by a 4×4 Latin Square [36]. Each task started with the home page of the website (depending on the website corresponding to the task), which consists of a list of popular health topics and a search input box. Participants were allowed to search, navigate, and browse the website freely. The only restriction was that they could not open and use websites other than the testing one. Screen captures, activity logs, and Web browsing histories were recorded for analysis.

At the halfway point and at the completion of the study session, a short semistructured interview was conducted, mainly for collecting verbal feedback about the tasks. This also gave the researcher a chance to collect feedback from the users about their experiences. The interview questions can be seen in Multimedia Appendix 3. Interviews were recorded and transcribed for future analysis.

Participant Recruitment

Participants were recruited via multiple channels, such as a University of Melbourne mailing list, electronic bulletin boards, and fliers posted in student lounges. We also used social media such as Facebook and Twitter to increase the exposure. We sought participants who were over aged 18 years and possessed previous experience of searching information on the Internet for comparing the 2 websites. Participants received no incentive to take part in the research.

Data Analysis

Multiple methodologies were used to analyze the various data obtained in this study. For qualitative data, semistructured interviews were transcribed and processed with content analysis [37,38]. Themes obtained from the analysis were used to investigate the influence on the design. In addition, screen recordings were reviewed and coded by the researcher [39,40]. This information was used to categorize and compare the patterns of health information–seeking behaviors in both the websites.

For quantitative analysis, we selected 4 metrics from previous research for measuring user interactions in the health information–seeking process [24,41,42]. Page reads and task duration reflect the amount of information accessed and the engagement with the website. Clicking on links (and tiles) indicates the effort of in-depth understanding about a topic and represents the depth of search. Query reformulation is an essential concept in exploratory search for measuring the degree of information exploration [43,44]. In the context of this study, query reformulation refers to issuing a new search query in the baseline website or adjusting the sliders in BHX.

Statistical calculations on the quantitative data were performed using R version 3.2.3. We applied Wilcoxon signed-rank test [45] to verify the statistical significance and compute the effect size between the baseline and BHX in each category (ie, focused and exploratory) of search tasks. This test does not require the normality of data [46].

Results

Participants

Overall, 31 participants took part in the study (N=31). Among the participants, 15 (48%) were male, and 16 (52%) were female. The average age was 33.9 (standard deviation=12.67, median=29) with the range from 20 to 72. Regarding the source of recruitment, 19 (61%) reported being students in the university; 10 (32%) were staff of the university; and 2 (7%) were recruited externally.

Health Information–Seeking Behaviors

Table 2 summarizes the annotations of health information–seeking behaviors (Figure 2) from the review of screen captures. For focused search tasks (Tasks A and B), most seeking behavior was consistent with Quick Fact Seeking (42% and 52%, respectively). However, more than half of participants demonstrated All-around Skimming (58%) in both the exploratory tasks (Tasks C and D). The results reinforce our previous results [19,27] that seekers adopt different information–seeking behaviors due to the different nature of search tasks.

Table 2. Seeking behaviors observed in our participants.

Seeking behavior	Focused search, n (%)		Exploratory search, n (%)	
	Task A (Baseline)	Task B (BHX[a])	Task C (Baseline)	Task D (BHX)
Quick Fact Seeking	13 (42)	16 (52)	4 (13)	3 (10)
Focused Reading	8 (26)	5 (16)	2 (6)	2 (6)
All-around Skimming	9 (29)	9 (29)	18 (58)	18 (58)
Knowledge Digging	1 (3)	1 (3)	7 (23)	8 (26)
Total	31 (100)	31 (100)	31 (100)	31 (100)

[a] BHX: Better Health Explorer

The figures display a distinct pattern in terms of the dimension of Research Tactics in our model of health information–seeking behaviors (Figure 2). Seekers demonstrated different behaviors across the task type, even with the same website. In focused search tasks, most participants (68% in both Tasks A and B) adopted basic Research Tactics (ie, Quick Fact Seeking and Focused Reading) behaviors. However, large portions of participants (81% in Task C; 84% in Task D) showed extensive Research Tactics (ie, All-around Skimming and Knowledge Digging) in exploratory search tasks. Nevertheless, all 4 types of information-seeking behaviors were observed in the study.

This reflects the diverse composition of health information seekers in the population of health website users.

User Interactions

Participants illustrated different levels of user interactions across focused and exploratory tasks. In focused search tasks (Tasks A and B), we did not observe significant differences between the baseline and BHX (Table 3). Participants read a similar number of pages, clicked on a similar number of links, and spent a similar amount of time in both the websites. Nevertheless, the number of query reformulations is substantially higher for BHX with a mean of 3.3 compared with 0.8 for the baseline (*P*<.001).

Table 3. User interaction figures in focused search tasks.

	Task A (Baseline)	Task B (BHX[a])	Wilcoxon signed-rank test		
	Mean (SD)	Mean (SD)	Z	*P*	r
Page reads	2.7 (1.5)	2.7 (1.9)	0.230	.847	.029
Task duration (seconds)	285 (115)	271 (142)	−0.598	.558	−.076
Clicks on links (baseline)/tiles (BHX)	2.6 (1.7)	2.5 (2.0)	−0.109	.927	−.014
Query reformulation	0.8 (0.8)	3.3 (3.5)	3.942	.000[b]	.501

[a] BHX: Better Health Explorer

[b] *P*<.001

However, the figures demonstrate a different pattern in exploratory search tasks (Table 4). BHX users presented a substantial higher number of pages read (*P*<.001) and more

query reformulations with a large effect size (*P*<.001). They also spent more time on exploring information (*P*=.034), and followed up more links in the website (*P*<.001).

XSL•FO

RenderX

Table 4. User interaction figures in exploratory search tasks.

	Task C (Baseline)	Task D (BHX[a])	Wilcoxon signed-rank test		
	Mean (SD)	Mean (SD)	Z	P	r
Page reads	3.9 (1.7)	5.7 (2.5)	3.986	.000[b]	.506
Task duration (seconds)	364 (153)	410 (154)	2.107	.034[c]	.268
Clicks on links (baseline) or clicks on tiles (BHX)	2.7 (1.9)	4.8 (2.4)	3.857	.000[b]	.490
Query reformulation	0.3 (0.6)	4.2 (3.3)	4.584	.000[b]	.582

[a]BHX: Better Health Explorer

[b]$P<.001$.

[c]$P<.05$.

The raw results of the quantitative figures can be obtained from Multimedia Appendix 4.

Qualitative Feedback

This subsection presents the qualitative feedback from the participants, interpreted together with the *in situ* observations of the primary researcher. After qualitative content analysis, we have derived 4 main themes for our design.

Dynamic Information Scope

Participants reported that our design helped them to seek information within a dynamic information scope. Most of the time they looked for the directly relevant information, but, in certain cases, they also needed to broaden to other topics. Participants reported that BHX was useful in such cases to discover topics. This was accomplished by adjusting the sliders to allow diverse topics to appear. In contrast, traditional health websites and search engines require issuing new search queries to achieve the same result. Our approach lowered the cognitive load of constructing new search queries and thus was perceived easier to use.

> *It (BHX) did help me find topics, and I think here I found topics that were related to what I was reading… In another one (the baseline website) I need to look for topics. This one (BHX) showed me topics. So it is easier.* [Participant #26]

> *People like to see the context, (and) this one (BHX) shows me the context.* [Participant #8]

> *This (BHX) is very good for discovering related information.* [Participant #14]

> *I expect things that are somehow relevant.* [Participant #6]

Moreover, participants appreciated that BHX showed diverse types of information around the topic, which provided opportunities for approaching the problem from different perspectives.

> *It (BHX) gives me more options about the topic.* [Participant #15]

> *By giving you choices rather than just coming up with the top things, I think it makes (the system) more interesting to use.* [Participant #3]

> *There are a range of things that are sort of related… and things are not much related. There are a range of different aspects. So I think that has a good diversity.* [Participant #3]

Serendipity and Curiosity

Serendipity, meaning that people are surprised by seeing valuable things that they have not thought of [47], was found critical in this study. Serendipitous findings happened when seekers did not know much about the health topics, and in such a case, BHX reminded them the existing information they did not consider.

> *Giving me different options that I did not consider.* [Participant # 15]

> *It is useful if I can get other information connecting (a sickness) to other subjects that I have not thought of before.* [Participant # 23]

> *I never thought about this and then (it showed up). Ah! This is good.* [Participant # 26]

Curiosity was also an important factor for engaging users in the information exploration process. The design of BHX was observed to stimulate curiosity in the health information–seeking process.

> *It (BHX) is engaging because you can kinda play with it and see what you get. You have a reasonable expectation what sort of things you are going to get and what exactly you are getting out of it. It is sort of curious.* [Participant # 13]

> *I am curious to see what it is all about.* [Participant # 18]

> *Because when you started reading something, the another one (tile) gave me more options—something that captured my eyes and it is interesting.* [Participant # 3]

Trust Issues

BHX offers a "fuzzy" approach for seeking health information that is very different from search engines, which always provide best matches to the search terms. In this study, we observed that participants had different opinions about the sometimes unexpected results displayed. The following section includes some positive feedback about this approach.

(When seeing something unexpected came up...) I trust your system! I do not necessarily think that means something is wrong. I think that means maybe there is a connection there (which) I was not aware of. It is more interesting rather than a problem. [Participant # 20]

This (an unexpected thing) is the information that I do not have in my mind, so it was put there like opportunities of knowledge. Because I was not aware, it gives me awareness of things that are related to the topic. Many times I do not know what to look for. I think it gives me awareness. [Participant # 15]

Other participants suggested that the unexpected results would be a problem and these endangered the trust between the user and the website. The primary goal of any design should offer the best results as possible.

That makes me worry. You start to bring up information which is perhaps gonna scaring people as well. It is about to get the right stuffs to the top first. [Participant # 10]

(The accuracy) It is difficult to judge for me. Unless... for example, if I am an expert of a particular topic, then I can actually make a correct call, whether this is giving me the correct information. [Participant # 23]

Delightful User Experience

We noticed that the design of BHX delivered a delightful experience to the users. The interactivity in the information exploration process gave them the feeling of gaining more control over the website and thus increased their satisfaction. Also, the opportunities for acquiring new knowledge led to a higher engagement between the participants and the website.

The interactive nature of it... You have more control over whatever information you get. [Participant #24]

Exploring using this website (BHX) is easier. You have more options. Whereas the other website (the baseline website) is fixed, (where) you do not have much to control. [Participant #5]

It is good because it makes me feel more knowledge after clicking here and finding something. [Participant #15]

I do enjoy using this (BHX). Because there are some interesting topics here. I am definitely getting interesting topics here. There are many many topics popping up which I am happy to follow up. There are so many I can choose from. [Participant #6]

Discussion

Principal Findings

In this study, we have identified the combinations of heterogeneous behaviors in the health information–seeking process, in terms of different behaviors and user interactions. Overall, our design improves the user experience of health information seeking and better supports the needs of health information seekers.

The results provide insights about our innovative design for health information seeking. Participants enjoyed using the interactive UI to find information. Such a design was considered to provide a more diverse and dynamic information scope. Meanwhile, the evaluation illustrates that a design that encourages serendipitous findings can provide hints for further search directions, when people have little prior knowledge about the health topic. Interactivity and having control over the system are also factors that engage with information seekers. In terms of quantitative analysis, BHX was shown to perform better in exploratory tasks across all metrics. This highlights the importance of directly supporting information exploration in health information seeking.

These findings imply that the design of BHX is superior in supporting the various health information–seeking behaviors. In the data analysis process, we have identified 4 design considerations that are crucial for health websites and health information–seeking apps, including: (1) providing a dynamic information scope; (2) supporting serendipity; (3) considering trust implications; and (4) enhancing interactivity. We argue that these aspects play a role in the design of consumer health websites and health information–seeking apps, for providing a better environment for seeking and conveying health information.

In the following subsections, we will further discuss these design considerations with the lessons learnt from this experiment.

Providing a Dynamic Information Scope

For both focused and exploratory search tasks, we have identified that seekers demand information from a dynamic information scope. Although seekers often prefer to retrieve direct and relevant information to the current context, they would like to see more diverse options (eg, information of other topics but still relevant) at the same time. This helps seekers to understand the health issues from different perspectives.

We propose that the variance is an outcome of the coexistence of both focused and exploratory search approaches in health information seeking [17,19] and related to the uncertainty that arises from the health issue [17,24,48]. When possessing only little knowledge about the health problem, seekers will have an unclear search target. As a result, they look for more different options for learning and setting up a clear direction to handle the scenario. After reading enough information, the search scope turns into a more focused one because the uncertainty is reduced. The same level of diversity is not required at the point when seekers have low uncertainty.

In addition, the nature of search tasks affects information needs and the search approach used. An ill-formed and widely open health scenario, such as a self-initiated search about certain symptoms, leads to a more exploratory search and thus requires a wider range of health information. However, a search after a diagnosis often leads to a focused search because of a clear, predefined search goal.

BHX uses a combination of sliders to facilitate exploration in a dynamic information scope. Both themes from the qualitative analysis (Dynamic Information Scope and Delightful User Experience) and quantitative results reflect the success of this

XSL•FO
RenderX

design component. Although many traditional health websites provide search functions, few allow exploratory search through offering a broad and dynamic range of information. To explore a new topic, users have to issue new search queries. In contrast, seekers using BHX can adjust the sliders and review the results that emerge with the new criteria. This process is more convenient, and the cognitive workload is lower compared with that while making a new search query.

Based on the aforementioned observations, we consider that access to a dynamic information scope is critical for health information seeking, as health information seekers demonstrate complicated needs and heterogeneous behaviors throughout their searches.

Supporting Serendipity

Serendipity, which refers to scenarios of discovering useful information with randomness and pleasure [49], can be a factor to assist in the information seeking process [25,50]. Some existing work has pointed out that health information seeking involves certain degree of serendipity [28,51], resulting in good search outcomes with less effort [52].

Similar conclusions are found in our experiment. Revealed in the qualitative themes, seekers demonstrate a positive attitude to serendipities. Seekers are surprised when they spot information that they did not anticipate on the screen and even show excitement if the information "hits" their needs. Although serendipity is often found in exploratory search [25,53], we argue that serendipity is relevant to both focused search and exploratory search because of the interleaving search approaches in health information seeking.

Serendipity associates with certain defining properties of health information seeking. First, general seekers are not experts in the health domain and do not have the knowledge to create proper search queries [6-8]. Observed in the study, they use the serendipitous findings as hints and directions for future searches. Also, serendipity is deemed as a way of making unforeseeable connections across different knowledge, and such connections are usually preferred, as people will favor using the least effort in finding information [54].

Curiosity is another factor to drive health information seeking, especially in the event of casual reading, for example, reading health information for self-learning rather than facing an immediate threat [17]. In this research, serendipity is found to trigger curiosity, and therefore, further readings are carried out. In many cases, casual seekers do not have a specific search target; hence, a fuzzy query approach (such as the design of BHX) can increase the chance of returning information that is interested to the user. This has been deemed as another manner approaching the information retrieval problem [55].

BHX gives an example of designing for serendipity. Instead of returning accurate search results as seen in most search engines, our design of providing a fuzzy search assists in health information exploration and enhances the user experience. With the advantages identified in this evaluation, we recommend to add serendipitous elements for health information seeking.

Considering Trust Implications

Trust is always an important component of health websites. Various guidelines have been proposed to ensure genuine and trustworthy online Web-based information [56,57]. Although these guidelines focus mainly on the quality and the accuracy of the content, the trust issues that arise from the information presentation and the UI are seldom discussed. In our study, we have identified new challenges for trustworthiness that result from the novel design.

By using a fuzzy approach for exploring information, it is inevitable that users will encounter some results that appear "weird" or "irrelevant." From our observations, we identify 2 causes underlying this issue: the seeker does not recognize the connection between the displayed information and the current context or the user does not fully understand the meaning of the information visualization of the UI.

We have identified 2 extreme responses when seekers observe the questionable results. Some users trust the website and believe that they do not possess the expertise to understand why these results appear. They realize that the "irrelevant" result is an opportunity to learn the unknown and leads to serendipitous findings. For this scenario, as seekers may not be able to judge the validity of the results, research should focus on the accuracy of the algorithm for providing truly relevant information, even for serendipities.

Other seekers assume that the display of doubtful information is a fault of the system. These seekers often possess higher health literacy and more knowledge about the health issue, so they believe that the system gives incorrect information. Eventually the trust between seekers and the system will cease. For this type of seeker, the solution is to increase the transparency of how the system works, by clearly explaining the reasons behind the display of certain items in the UI. For example, the design can adopt different color codes or legends to indicate the results suggested by different heuristics or at different levels of confidence. Visual feedback (eg, highlights and animations) can also be added to represent the results originating from certain user activities.

Regardless of the circumstances, designers need to be careful in presenting health information when using new interaction techniques. User may have unexpected interpretations of the information presentation and the UI, as compared to what the designer expects. A comprehensive understanding of the users and usability tests can help to resolve such problems before release.

Enhancing Interactivity

This study shows that our design provides higher interactivity in comparison with traditional consumer health websites and therefore leads to higher user engagement and better user experience. This interactive nature gave users more options to retrieve more diverse information. Participants could manipulate this exploration through the UI, observe the changes to the results, and obtain new knowledge from this process. In contrast, the traditional experience relying on keyword search was seen as less playful and enjoyable. The positive findings are consistent with those of other research studies, which suggests

that interactivity has a positive impact on information seeking [10,58,59]. This highlights the potentiality for enhancing the interactivity of health information websites.

In addition, interactivity and engagement may be related to better performance and higher rates of returning to a health website. According to Flow research, a person who enters a mental state of complete engagement and immersion into an activity will have higher success rate and will have a greater chance of reusing the system [60,61]. In the BHX study, participants were observed to engage with the elements supporting interactive exploration of health information. Therefore, we expect that a design with better interactivity will bring a positive outcome in seeking and learning consumer health information. Future research might focus on revealing the relationship among interactivity, engagement, and user revisits.

On the basis of our experience of this research, there are challenges to implement a website with these promising features. Introducing an interactive experience often requires a new UI. This may not be easily accepted by users and may have negative implications for trust as discussed previously. The existing content may not be directly usable in the new UI and would need to be manipulated to fit the distinct underlying model. For example, in our website, articles were processed with a computer program to generate the metadata for the sliders and had to be reviewed manually. Such additional work may be very time consuming and require significant effort. Hence, we should be aware of these challenges despite the better outcomes achieved by the proposed design.

Limitations

Although most of our statistical tests are significant, the sample size is relatively small as compared with that of other quantitative studies. To reduce the impact, we applied a mixed approach to analyze both qualitative and quantitative data. In addition, the composition of the participants is mainly university members, which may not represent the general population of health information seekers. Future research will focus on a larger cohort of participants with a more diverse background.

Second, we evaluated all participants as a single group of health information seekers and did not capture attributes such as occupations and education level. For future research, studies focusing on different user groups may discover additional findings distinct to these groups.

Finally, only 1 website (ie, BHC) was chosen as the baseline, which limits the comparison only to the particular design of that website and therefore affects the generality of the results. Future work may evaluate the design implications in the context of other health websites or health information sources.

Conclusions

In this paper, we present and evaluate BHX, which is a novel interface design that aims to address several problems faced by users in health information seeking. Through an observational study, we discover that the design of BHX can improve findability and discovery of information, as well as enhance the overall user experience. Moreover, the study shows that a mix of health information–seeking behaviors needs to be handled by health websites, highlighting the importance of providing specific support for these behaviors.

The positive results of this study reflect the importance of understanding different health information–seeking behaviors, as well as designing to accommodate these behaviors. Although previous eHealth research has suggested designing for users and their needs [5,22,23], this research takes a further step and proposes designing for behaviors, for example, reading, researching, and exploring. In this regard, 4 design considerations are emerged from our research. These considerations will lead to better support of the heterogeneous and shifting behaviors of health information seekers and ease the process of obtaining Web-based health information for users. Therefore, these elements should be applied in future designs and in HCI research.

Although promising outcomes are observed in this study, our future work includes addressing the critical feedback about the information presentation, extending the information sources to include other health databases, and investigating the effects of particular design features (eg, sliders and the exploration UI) in the health information–seeking processes.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Demonstration video of Better Health Explorer.

[MOV File, 62MB - jmir_v18i6e145_app1.mov]

XSL•FO
RenderX

Multimedia Appendix 2

Task assignments of the participants.

[XLSX File (Microsoft Excel File), 10KB - jmir_v18i6e145_app2.xlsx]

Multimedia Appendix 3

Questions for guiding the semistructured interview.

[PDF File (Adobe PDF File), 23KB - jmir_v18i6e145_app3.pdf]

Multimedia Appendix 4

Demographic and user interaction figures (N=31).

[XLSX File (Microsoft Excel File), 13KB - jmir_v18i6e145_app4.xlsx]

References

1. Fox S, Duggan M. Pew Research Center's Internet & American Life Project. 2013. Health Online 2013 URL: http://pewinternet.org/Reports/2013/Health-online.aspx [WebCite Cache ID 6fD2xaEjp]
2. Eysenbach G. Design and Evaluation of Consumer Health Information Web Sites. Consum Heal Informatics 2005;34:60. [doi: 10.1007/0-387-27652-1_4]
3. Lorence DP, Greenberg L. The zeitgeist of online health search. Implications for a consumer-centric health system. J Gen Intern Med 2006 Feb;21(2):134-139 [FREE Full text] [doi: 10.1111/j.1525-1497.2005.0295.x] [Medline: 16336621]
4. Samuel H, Zaiane O, Zaiane J. Findability in Health Information Websites. In: International Conference on Biomedical and Health Informatics (BHI 2012). 2012 Presented at: International Conference on Biomedical and Health Informatics (BHI 2012); 2012; Hong Kong and Shenzhen, China p. 709-712. [doi: 10.1109/BHI.2012.6211681]
5. Lee K, Hoti K, Hughes JD, Emmerton LM. Consumer Use of "Dr Google": A Survey on Health Information-Seeking Behaviors and Navigational Needs. J Med Internet Res 2015;17(12):e288 [FREE Full text] [doi: 10.2196/jmir.4345] [Medline: 26715363]
6. Chapman K, Abraham C, Jenkins V, Fallowfield L. Lay understanding of terms used in cancer consultations. Psychooncology 2003 Sep;12(6):557-566. [doi: 10.1002/pon.673] [Medline: 12923796]
7. Keselman A, Browne AC, Kaufman DR. Consumer health information seeking as hypothesis testing. J Am Med Inform Assoc 2008;15(4):484-495 [FREE Full text] [doi: 10.1197/jamia.M2449] [Medline: 18436912]
8. Luo G, Tang C, Yang H, Wei X. MedSearch: a specialized search engine for medical information retrieval. 2008 Presented at: Proceedings of the 17th ACM Conference on Information and Knowledge Management (CIKM '08); 2008; Napa Valley, CA. [doi: 10.1145/1458082.1458104]
9. Bickmore TW, Utami D, Matsuyama R, Paasche-Orlow MK. Improving Access to Online Health Information With Conversational Agents: A Randomized Controlled Experiment. J Med Internet Res 2016;18(1):e1 [FREE Full text] [doi: 10.2196/jmir.5239] [Medline: 26728964]
10. Cline RJ, Haynes KM. Consumer health information seeking on the Internet: the state of the art. Health Educ Res 2001 Dec;16(6):671-692 [FREE Full text] [Medline: 11780707]
11. Birru MS, Monaco VM, Charles L, Drew H, Njie V, Bierria T, et al. Internet usage by low-literacy adults seeking health information: an observational analysis. J Med Internet Res 2004 Sep 3;6(3):e25 [FREE Full text] [doi: 10.2196/jmir.6.3.e25] [Medline: 15471751]
12. Lam MK, Lam LT. Health information-seeking behaviour on the Internet and health literacy among older Australians. electronic Journal of Health Informatics 2012;7(2):1.
13. Agree E, King A, Castro C, Wiley A, Borzekowski Dina L G. "It's got to be on this page": age and cognitive style in a study of online health information seeking. J Med Internet Res 2015;17(3):e79 [FREE Full text] [doi: 10.2196/jmir.3352] [Medline: 25831483]
14. Goldberg L, Lide B, Lowry S, Massett HA, O'Connell T, Preece J, et al. Usability and accessibility in consumer health informatics current trends and future challenges. Am J Prev Med 2011 May;40(5 Suppl 2):S187-S197. [doi: 10.1016/j.amepre.2011.01.009] [Medline: 21521594]
15. Haak M van den, Hooijdonk CV. Evaluating Consumer Health Information Websites: The Importance of Collecting Observational, User-driven Data. In: 2010 IEEE International Professional Communication Conference (IPCC). Enschede, Netherlands: IEEE; 2010 Presented at: 2010 IEEE International Professional Communication Conference (IPCC); 2010; NA p. 333-338. [doi: 10.1109/IPCC.2010.5530031]
16. Lee K, Hoti K, Hughes JD, Emmerton L. Dr Google and the consumer: a qualitative study exploring the navigational needs and online health information-seeking behaviors of consumers with chronic health conditions. J Med Internet Res 2014;16(12):e262 [FREE Full text] [doi: 10.2196/jmir.3706] [Medline: 25470306]

17.   Pang P, Chang S, Pearce J, Verspoor K. Online Health Information Seeking Behaviour: Understanding Different Search Approaches. In: Proceedings of the 18th Pacific Asia Conference on Information Systems (PACIS 2014).: AIS; 2014 Presented at: PACIS 2014; 2014; Chengdu, China URL: http://aisel.aisnet.org/pacis2014/229 [WebCite Cache ID 6hp4O96fB]

18.   Pang P, Verspoor K, Pearce J, Chang S. Better Health Explorer: Designing for Health Information Seekers. In: Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction (OzCHI'15). New York, New York, USA: ACM Press; 2015 Presented at: OzCHI 2015; Dec 2015; Melbourne, Australia p. 588-597 URL: http://dl.acm.org/citation.cfm?doid=2838739.2838772 [WebCite Cache ID 6hp8g3Jws] [doi: 10.1145/2838739.2838772]

19.   Pang P, Verspoor K, Chang S, Pearce J. Conceptualising health information seeking behaviours and exploratory search: result of a qualitative study. Health Technol 2015 Feb 10;5(1):45-55. [doi: 10.1007/s12553-015-0096-0]

20.   White R, Horvitz E. Cyberchondria. ACM Trans. Inf. Syst 2009 Nov 01;27(4):1-37. [doi: 10.1145/1629096.1629101]

21.   Cartright M, White R, Horvitz E. Intentions and Attention in Exploratory Health Search. In: The 34th Annual ACM SIGIR Conference (SIGIR'11).: ACM; 2011 Presented at: The 34th Annual ACM SIGIR Conference (SIGIR'11); 2011; Beijing, China. [doi: 10.1145/2009916.2009929]

22.   van Gemert-Pijnen Julia E W C, Nijland N, van LM, Ossebaard H, Kelders S, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]

23.   Alzougool B, Chang S, Gray K. The nature and constitution of informal carers' information needs: what you don't know you need is as important as what you want to know. information Research. 2013. (1) p. 1 URL: http://informationr.net/ir/18-1/paper563.html [WebCite Cache ID 6hp8ydBSh]

24.   White R, Roth R. Exploratory Search: Beyond the Query-Response Paradigm. In: Synthesis Lectures on Information Concepts, Retrieval, and Services. San Rafael, CA: Morgan and Claypool; Jan 2009:1-98.

25.   Marchionini G. Exploratory search. Commun. ACM 2006 Apr 01;49(4):41. [doi: 10.1145/1121949.1121979]

26.   Pearce J, Chang S, Alzougool B, Kennedy G, Ainley M, Rodrigues S. Search or Explore: Do you know what you're looking for? In: 2011 Australian Computer-Human Interaction Conference (OzCHI'11). 2011 Presented at: OzCHI'11; 2011; Canberra, Australia p. 253-256. [doi: 10.1145/2071536.2071576]

27.   Pang P, Verspoor K, Chang S, Pearce J. Designing for Health Exploratory Seeking Behaviour. In: Medical Information Retrieval (MedIR) Workshop in the 37th Annual International ACM SIGIR Conference (SIGIR 2014). 2014 Presented at: Medical Information Retrieval (MedIR) Workshop in the 37th Annual International ACM SIGIR Conference (SIGIR 2014); 2014; Gold Coast, Australia URL: http://ceur-ws.org/Vol-1276/MedIR-SIGIR2014-01.pdf [WebCite Cache ID 6hp93HQLa]

28.   Johnson J, Case D. Health information seeking. New York, NY: Peter Lang; 2012.

29.   Department of Health and Human Services. Better Health Channel URL: https://www.betterhealth.vic.gov.au [accessed 2016-02-11] [WebCite Cache ID 6fD389k3o]

30.   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]

31.   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]

32.   Buhi ER, Daley EM, Fuhrmann HJ, Smith SA. An observational study of how young people search for online sexual health information. J Am Coll Health 2009;58(2):101-111. [doi: 10.1080/07448480903221236] [Medline: 19892646]

33.   Kules B, Capra R, Banta M, Sierra T. What Do Exploratory Searchers Look at in a Faceted Search Interface? In: Proceedings of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries.: ACM; 2009 Presented at: Proceedings of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries; 2009; New York, NY, USA p. 313-322. [doi: 10.1145/1555400.1555452]

34.   Wildemuth B, Freund L. Assigning Search Tasks Designed to Elicit Exploratory Search Behaviors. In: Proceedings of the Symposium on Human-Computer Interaction and Information Retrieval.: ACM; 2012 Presented at: Proceedings of the Symposium on Human-Computer Interaction and Information Retrieval; 2012; New York, NY, USA p. 1-10. [doi: 10.1145/2391224.2391228]

35.   Hendahewa C, Shah C. Implicit search feature based approach to assist users in exploratory search tasks. Information Processing & Management 2015 Sep;51(5):643-661 [FREE Full text] [doi: 10.1016/j.ipm.2015.06.004]

36.   Williams E. Experimental Designs Balanced for the Estimation of Residual Effects of Treatments. Aust. J. Chem 1949;2(2):149. [doi: 10.1071/CH9490149]

37.   Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]

38.   Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

39.   Creswell J. Qualitative Inquiry and Research Design: Choosing Among Five Approaches. Thousand Oaks, CA: Sage Publications; 2013.

40. Creswell J. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Thousand Oaks, CA: Sage Publications; 2014.

41. Hodkinson C, Kiel G, Mccoll-Kennedy JR. Consumer web search behaviour: diagrammatic illustration of wayfinding on the web. International Journal of Human-Computer Studies 2000 May;52(5):805-830. [doi: 10.1006/ijhc.1999.0357]

42. Huang P, Lurie NH, Mitra S. Searching for Experience on the Web: An Empirical Examination of Consumer Behavior for Search and Experience Goods. Journal of Marketing 2009 Mar;73(2):55-69. [doi: 10.1509/jmkg.73.2.55]

43. Klouche K, Ruotsalo T, Cabral D, Andolina S, Bellucci A, Jacucci G. Designing for Exploratory Search on Touch Devices. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. USA: ACM; 2015 Presented at: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems; 18-23 Apr 2015; Seoul, Korea p. 4189-4198. [doi: 10.1145/2702123.2702489]

44. Zhang Y, Peiling W, Heaton A, Winkler H. Health Information Searching Behavior in MedlinePlus and the Impact of Tasks. In: 2012 ACM SIGHIT International Health Informatics Symposium (IHI 2012). 2012 Presented at: 2012 ACM SIGHIT International Health Informatics Symposium (IHI 2012); 2012; Miami, Florida, USA p. 641-650.

45. Wilcoxon F. Individual comparisons by ranking methods. Biometrics bulletin 1945:80-83.

46. McDonald J. Handbook of biological statistics. 3rd edition. Baltimore, Maryland: Sparky House Publishing; 2014.

47. Apted E, Choo C. General purposive browsing. Library Association Record 1971;73(12):228-230.

48. Wilson T. Information behaviour: An interdisciplinary perspective. Information Processing & Management 1997 Jul;33(4):551-572. [doi: 10.1016/S0306-4573(97)00028-9]

49. Thudt A, Hinrichs U, Carpendale S. The bohemian bookshelf: supporting serendipitous book discoveries through information visualization. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12). Austin, Texas, USA: ACM; 2012 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12); 2012; Austin, Texas, USA p. 1461-1470. [doi: 10.1145/2207676.2208607]

50. Foster A, Ford N. Serendipity and information seeking: an empirical study. Journal of Documentation 2003 Jun;59(3):321-340. [doi: 10.1108/00220410310472518]

51. Case D. Looking for Information: A Survey of Research on Information Seeking, Needs, and Behavior. 3rd ed. The Earth: Bingleymerald Group Publishing Limited; 2012.

52. David Johnson J. Health-related information seeking: Is it worth it? Information Processing & Management 2014 Sep;50(5):708-717 [FREE Full text] [doi: 10.1016/j.ipm.2014.06.001]

53. White RW, Kules B, Drucker SM, schraefel M. Introduction. Commun. ACM 2006 Apr 01;49(4):36. [doi: 10.1145/1121949.1121978]

54. Case D. Principle of least effort. In: Theories of Information Behavior. USA: Information Today, Inc; 2005.

55. Toms E. DELOS Work Inf Seek Search Querying Digit Libr Internet. 2000. Serendipitous Information Retrieval URL: http://www.ercim.eu/publication/ws-proceedings/DelNoe01/3_Toms.pdf [accessed 2016-05-20] [WebCite Cache ID 6heELjlAl]

56. Boyer C, Selby M, Scherrer J, Appel R. The Health On the Net Code of Conduct for medical and health Websites. Computers in Biology and Medicine 1998 Sep;28(5):603-610 [FREE Full text] [doi: 10.1016/S0010-4825(98)00037-7]

57. Commission of the European Communities, Brussels. eEurope 2002: Quality Criteria for Health Related Websites. J Med Internet Res 2002 Dec;4(3):E15 [FREE Full text] [doi: 10.2196/jmir.4.3.e15] [Medline: 12554546]

58. Niu X, Hemminger BM, Lown C, Adams S, Brown C, Level A, et al. National study of information seeking behavior of academic researchers in the United States. J. Am. Soc. Inf. Sci 2010 Feb 24;61(5):869-890. [doi: 10.1002/asi.21307]

59. Crutzen R, Cyr D, de Vries Nanne K. The role of user control in adherence to and knowledge gained from a website: randomized comparison between a tunneled version and a freedom-of-choice version. J Med Internet Res 2012;14(2):e45 [FREE Full text] [doi: 10.2196/jmir.1922] [Medline: 22532074]

60. Csikszentmihalyi M. Beyond Boredom and Anxiety. San Francisco: Jossey-Bass Publishers; 1975.

61. Pearce J, Murphy J, Patman D. Using a Scenario-planning Tool to Support an Engaging Online User Experience. In: Proceedings of the 18th Australia Conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments. USA: ACM; 2006 Presented at: Proceedings of the 18th Australia Conference on Computer-Human Interaction; 2006; Australia p. 167-174. [doi: 10.1145/1228175.1228206]

Abbreviations

BHC: Better Health Channel
BHX: Better Health Explorer
HCI: Human-computer Interaction
UI: User interface

XSL•FO

RenderX

<u>Original Paper</u>

Evaluation of Chlamydia Partner Notification Practices and Use of the "Let Them Know" Website by Family Planning Clinicians in Australia: Cross-Sectional Study

Rebecca J Guy[1], PhD; Joanne M Micallef[1], PhD; Julie Mooney-Somers[2], PhD; Muhammad S Jamil[1], MPH, MHM, MBBS; Caroline Harvey[3], MPH, MPM, MBBS, FRACGP, DRANZCOG; Deborah Bateson[4,5], MBBS, MSc, MA (Oxon); Caroline van Gemert[6], MPH, MAppEpid; Handan Wand[1], PhD; John Kaldor[1], PhD

[1]The Kirby Institute, UNSW Australia, Sydney, Australia

[2]Centre for Values, Ethics and the Law in Medicine, University of Sydney, Sydney, Australia

[3]Institute for Urban Indigenous Health, Brisbane, Australia

[4]Family Planning New South Wales, Sydney, Australia

[5]Discipline of Obstetrics, Gynaecology and Neonatology, The University of Sydney, Sydney, Australia

[6]Centre for Population Health, Burnet Institute, Melbourne, Australia

Corresponding Author:
Rebecca J Guy, PhD
The Kirby Institute, UNSW Australia
Wallace Wurth Building
Sydney, 2052
Australia
Phone: 61 2 9385 0978
Fax: 61 2 9385 0920
Email: rguy@kirby.unsw.edu.au

Abstract

Background: Chlamydia, caused by *Chlamydia trachomatis*, is the most common reportable infection in many developed countries. Testing, treatment, and partner notification (PN) are key strategies for chlamydia control. In 2008 the Let Them Know (LTK) PN website was established, which provided means for people to send anonymous PN messages by text messaging (short message service, SMS), email, or letter.

Objective: We evaluated PN practices among Australian family planning clinicians following chlamydia diagnosis and assessed how often clinicians refer their patients to the LTK website.

Methods: A mixed methods approach included a Web-based cross-sectional survey of Australian family planning clinicians to examine PN attitudes and practices and focus groups to explore the context of LTK website use.

Results: Between May 2012 and June 2012, all clinicians from 29 different family planning services (n=212) were invited to complete the survey, and 164 participated (response rate=77.4%); of the clinicians, 96.3% (158/164) were females, 56.1% (92/164) nurses, and 43.9% (72/164) doctors. More than half (62.2%, 92/148) agreed that PN was primarily the client's responsibility; however, 93.2% (138/148) agreed it was the clinician's responsibility to support the client in informing their partners by providing information or access to resources. Almost half (49.4%, 76/154) of the clinicians said that they always or usually referred clients to the LTK website, with variation across clinics in Australian states and territories (0%-77%). Eleven focus groups among 70 clinicians at 11 family planning services found that the LTK website had been integrated into routine practice; that it was particularly useful for clients who found it difficult to contact partners; and that the LTK letters and fact sheets were useful. However, many clinicians were not aware of the website and noted a lack of internal clinic training about LTK.

Conclusions: The LTK website has become an important PN tool for family planning clinicians. The variation in referral of patients to the LTK website and lack of awareness among some clinicians suggest further promotion of the website, PN training, and clinic protocols are warranted.

XSL•FO
RenderX

KEYWORDS

Chlamydia trachomatis; partner notification; Internet

Introduction

Chlamydia, caused by *Chlamydia trachomatis*, is the most common reportable infection in the United States, Australia, and European countries [1-3]. In 2014 more than 1.4 million new diagnoses were reported in the United States [1], and 86,000 chlamydia cases were notified in Australia [3]. However, more than three-quarters (76%) of infections remain undiagnosed at any point in time [4]. Testing, treatment, and partner notification (PN) are key strategies for chlamydia control. Partner notification and testing has been shown to reduce reinfection rates in index cases [5]. Mathematical modeling suggests that, in a population-wide screening program, the treatment of current partner is the most effective strategy for preventing reinfection of index cases and reducing further chlamydia transmission at the population level [6].

Clinical guidelines in the United Kingdom and Australia recommend testing and treating all sexual partners in the last 6 months [7,8], whereas US guidelines recommend to treat all partners in the last 2 months or the most recent partner if the last sexual contact was more than 2 months ago [9]. Despite clinicians recognizing its importance, PN has long been recognized as a challenge for clinicians and patients alike, because of the sensitivities involved in disclosing and informing [10]. Most clinicians report they would like additional supportive resources, including websites [10]. New, accessible, and convenient approaches are needed to inform partners of their potential disease exposure.

In December 2008 the Let Them Know (LTK) website was launched in Australia and provided means for people to send anonymous PN messages by short message service (SMS) text messaging, email, or letter (Figure 1) [11]. The LTK website was the first in the world to enable young people to notify their partners anonymously using SMS text messaging, whereas other systems only offered electronic postcards or email. The LTK website also includes fact sheets for sexually transmitted infections (STIs) and letter templates with testing and treatment recommendations for the partners to pass on to their doctors. The website was developed by the Melbourne Sexual Health Centre in Victoria with some information on the website (contact details and letters) specific to this clinic; it was later adapted

for use in New South Wales and then Australia-wide by customization to local resources [11]. The website was developed originally for chlamydia and later adapted for other STIs [12]. No specific promotion of the website occurred across Australia.

In the past few years, other Internet-based PN services have been developed [13-18], and websites have also been used to promote chlamydia screening in young people and providers [19,20]. The "WhyTest" service for gay men in Australia included SMS text messaging and email notification [13]; the "inSPOT" service in the United States enabled notification by postcards and emails [16] and has since been replicated elsewhere [17,18]; and in the Netherlands, the "suggest a test" service offered SMS text messaging, email, postal letter, or a personal message to notify sexual contacts [14]. Evaluation of these PN websites has mainly focused on website usage and showed the SMS text messaging function is far more popular than email [13,14].

Because of the inherent nature of these services, demonstration of effectiveness is very challenging. To our knowledge only one randomized controlled trial has evaluated the impact of PN websites (inSPOT) on partner treatment among men who have sex with men [15]; however, the website only provided email and postcard services, which have been shown to be less popular than SMS text messaging. A previous evaluation of the LTK website has shown that people who used the service reported they were more likely to contact a partner because of the website [12], which should ultimately lead to a greater uptake of PN overall. However, studies in the United States show that awareness and uptake among the target group is low [17,18]. It is possible that the low uptake is due to clinicians not promoting the services actively to their clients.

We evaluated PN practices among family planning clinicians following chlamydia diagnosis and assessed how often clinicians refer patients to the LTK website to notify their partners. Although the LTK website offered PN for other STIs, we focused on chlamydia as the prevalence in young people in Australia is higher compared with other STIs including gonorrhea [3]. The study was conducted in the context of a broader study assessing chlamydia testing and management practices at Australian family planning clinics.

Figure 1. Let Them Know website interface screenshot.

Methods

Setting

The study was undertaken among clinicians, both doctors and nurses, working in Australian family planning clinics, which provide sexual and reproductive health (SRH) services and are located in all Australian states and territories. These clinics have a high caseload of young people aged 16-29 years who are sexually active and at risk of chlamydia infection, with more than 90% females [21]. The clinics are run by independent, nongovernment, not-for-profit organizations responsible to a voluntary board of directors.

Nurses in family planning clinics work in various roles within and between states. Most nurses have SRH qualifications but the scope of practice includes specialized SRH advanced practice nurses, which may include limited medication supply, various autonomous clinical consulting roles, and phone advice or clinical practice support. State or territory legislation governs nurse medication supply, which affects the degree of autonomy in treating chlamydia. For example, in some states most family planning nurses are authorized under legislated drug therapy protocols to autonomously supply treatment for people diagnosed with chlamydia and their contacts and would, as part of their duty of care, provide information on and support for PN. Whereas nurses in other jurisdictions may need to refer all chlamydia cases to family planning doctors for management and would generally not be involved in PN or support of clients.

Study Design

We used a mixed methods design involving a cross-sectional survey to examine chlamydia testing and management, and PN attitudes and practices among clinicians, and focus groups to explore the context of chlamydia management and PN.

Cross-Sectional Survey

All doctors and nurses recorded as being employed clinically at all 29 Australian family planning clinics (as of April 2012) were invited to complete a survey by email. Researchers sent an introductory email to the clinic manager or administrative officer, who circulated it among clinic staff members. The email contained the link to the Web-based survey. The clinic manager also provided the number of individual staff working at each clinic to calculate the response rate. Posters were also displayed in staff areas to raise awareness of the survey. After 4 weeks a reminder was sent to the clinic representative to encourage nonresponders to complete the survey.

The Web-based survey captured clinician demographics, experience, attitudes, and practices related to chlamydia testing and management (retesting and PN). Most questions sought responses on a 5-point Likert scale from strongly agree to strongly disagree or a 4-point scale of always, usually, sometimes, or never. The questionnaire was tested with clinicians for its content, language, and feasibility of questionnaire length.

Descriptive statistics were used to examine the responses to survey questions about chlamydia PN attitudes and practices.

Focus Groups

The cross-sectional survey and focus groups were conducted sequentially, not in parallel. We conducted surveys before the focus groups to allow us to use survey findings to purposively sample clinics [22]. To ensure a diverse sample of clinics we selected across a range of urban and regional clinics across Australia, client demographics (eg, youth focused), reported PN strategies, reported retesting strategies, and, finally, reported responses to screening in an asymptomatic client scenario. In selected clinics, the clinic representative invited all doctors and nurses to attend. Some clinics also invited health promotion

officers. Focus groups were conducted over lunchtime or directly after clinic hours. Each group was facilitated by 2 researchers and audio recorded.

Focus groups began with a general discussion about the clinical setting, client population, and the local ethos around chlamydia prevention. Then in relation to (1) chlamydia testing, (2) PN and treatment (including the use of technology such as LTK), (3) retesting, they were asked to discuss an example of practice where things went well, how they know when things are working well, and how they think things could work well more often. Focus groups were transcribed and data analyzed using thematic analysis [23]. NVivo 10.0 was used to support analysis.

Ethical Approval

Ethical approval was obtained from Family Planning NSW Ethics Committee and Family Planning Victoria Human Research Ethics Committee.

Results

Participants

Between May 2012 and June 2012, all clinicians employed by Australian family planning clinics (n=212) were invited to complete the Web-based survey and 164 participated, giving a response rate of 77.4%; of the clinicians, 96.3% (158/164) were females, 56.1% (92/164) were nurses, and 43.9% (72/164) were doctors. All 29 Australian family planning clinics were represented in the survey. More than half (56.1%, 92/164) of the participants were aged 45 years or older, 47.0% (77/164) had worked at the current organization for 6 years or more, 69.5% (114/164) as a clinician with a special interest in reproductive and sexual health, and 42.7% (70/164) worked at a family planning clinic less than 10 hours per week. About a third of the clinicians (31.7%, 52/164) managed more than 3 female clients with a positive chlamydia test result per month, 45.7% (75/164) saw 1-3 per month, and 20.1% (33/164) less than 1 per month (Table 1). Focus groups were held in 11

metropolitan and regional clinics across Australia and involved a total of 70 nurses, doctors, or health promotion officers (range 4-11 participants per clinic).

Partner Notification Attitudes

Almost all clinicians (99.3%, 147/148) strongly agreed or agreed that PN is an important strategy for preventing reinfection, and 97.3% (144/148) strongly agreed or agreed it is an important public health strategy to reduce the community prevalence of chlamydia. More than half (66.2%, 98/148) of the clinicians strongly agreed or agreed that PN is difficult as clients do not always feel comfortable talking to partners about chlamydia, and 47.3% (70/148) found it difficult because clients do not like to name their partners or are unable to name them (eg, they did not know the partner's name or contact details). On the other hand, relatively few (12.2%, 18/148) strongly agreed or agreed that PN is too difficult to implement (Figure 2).

More than half (62.2%, 92/148) of the participants strongly agreed or agreed that PN was primarily the client's responsibility, whereas 10.1% (15/148) strongly agreed or agreed it was primarily the clinician's responsibility to notify the partner. Nevertheless, the vast majority (93.2%, 138/148) strongly agreed or agreed it was the clinician's responsibility to support the client in informing their partners by providing information or access to resources. Focus group strongly reflected these findings; for example:

> *We teach that you're responsible for making sure they're aware that their partners should be contacted and treated, and that the clinician is happy to help that process. Whether that's through the Let Them Know website, or whether they actually want the clinician to do it themselves, which would be pretty unusual, I think. But the responsibility is to make sure the patient's aware that the partner should be notified and treated, and tested. So there's no - I don't think any of the clinicians probably feel that it's their responsibility to contact the partners.*

Table 1. Description of Australian family planning clinician survey participants.

Characteristics	Sub-category	n (%)
Provider type	Nurse	92 (56.1)
	Doctor	72 (43.9)
Sex	Female	158 (96.3)
	Male	3 (1.8)
	Missing	3 (1.8)
Age, years	<35	26 (15.9)
	35-44	45 (27.4)
	45-54	58 (35.4)
	55+	34 (20.7)
	Missing	1 (0.6)
Clinic location	Urban	103 (62.8)
	Regional/remote	61 (37.2)
Years worked at current organization	<1	22 (13.4)
	1-3	32 (19.5)
	>3-5	31 (18.9)
	>5	77 (47.0)
	Missing	2 (1.2)
Years worked as a clinician with a special interest in sexual/reproductive health	<3	25 (15.2)
	3-5	24 (14.6)
	>5	114 (69.5)
	Missing	1 (0.6)
Hours per week providing clinical services in a general family planning service	<10	70 (42.7)
	10-20	61 (37.2)
	21-30	29 (17.7)
	≥31	3 (1.8)
	Missing	1 (0.6)
Approximate chlamydia test requested per week	None	2 (1.2)
	1-5	80 (48.8)
	6-10	50 (30.5)
	11-20	25 (15.2)
	>20	5 (3.0)
	Missing	2 (1.2)
Female clients with a positive chlamydia test managed	None	11 (6.7)
	<1 per month	22 (13.4)
	1-3 per month	75 (45.7)
	1-2 per week	36 (22.0)
	3-5 per week	14 (8.5)
	≥6 per week	2 (1.2)
	Missing	4 (2.4)
Male clients with a positive chlamydia test managed	None	34 (20.7)
	<1 per month	70 (42.7)

Characteristics	Sub-category	n (%)
	1-3 per month 43 (26.2)	
	1-2 per week	12 (7.3)
	3-5 per week	2 (1.2)
	≥6 per week	0
	Missing	3 (1.8)

Figure 2. Partner notification attitudes among Australian family planning clinicians (n=148).
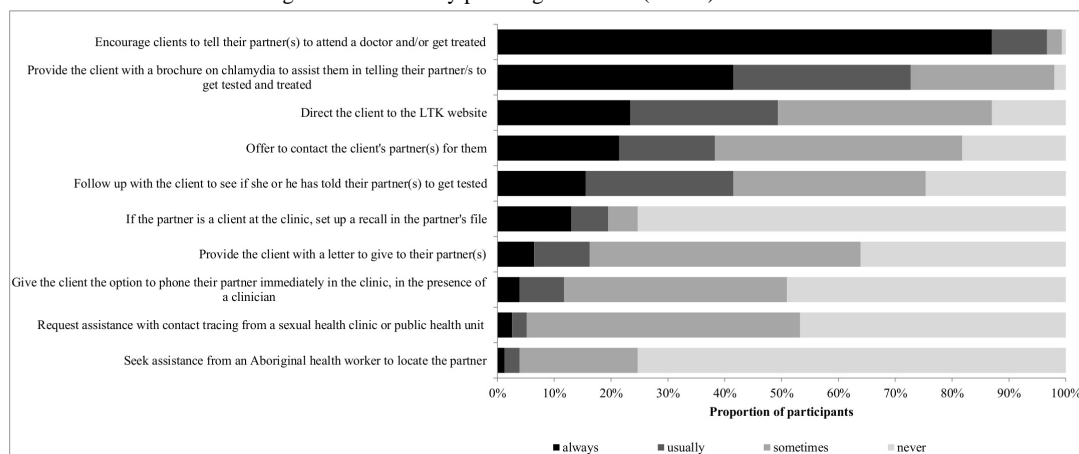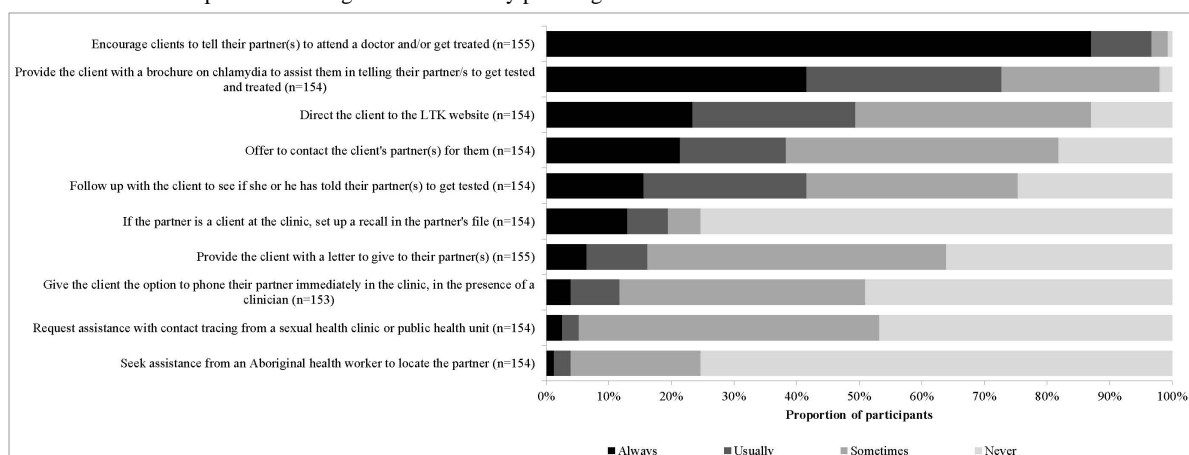

Partner Notification Practices

The most frequently reported strategy for PN was always or usually encouraging clients to undertake responsibility for communicating with sexual partners (96.8%, 150/155), followed by providing the clients with a brochure about chlamydia (72.7% 112/154), directing the clients to the LTK website (49.4%, 76/154), monitoring the clients to confirm they had notified their partner (41.6%, 64/154), and offering to contact the client's partners (38.3%, 59/154; Figure 3). The majority of clinicians stated, consistent with Australian clinical guidelines [7,24], that they encouraged contacting all sexual partners in the past 6 months before the diagnosis (77.1%, 118/153), with 11.1%

(14/153) stating 3 months or less, and 11.8% (18/153) stating 1 year or more. Focus groups revealed family planning clinics did not have specific protocols around the practice of PN; instead, individual clinicians tailored their approach to client need:

> *We have a policy around contact tracing, that we do it, but how it's done depends on that individual client, really. It's basically based on whether they can notify their partners. And if they can't or they don't want to, then we will do that but we don't actually have a whole protocol for notifying patients and partners and stuff, do we?*

Figure 3. Partner notification practices among Australian family planning clinicians.

Partner Notification Barriers

Three key challenges to PN emerged in the focus groups. First, across focus groups there was uncertainty about what was expected of clinicians in relation to PN. As indicated in the survey findings, there was consensus on the importance of telling clients they should inform their contacts. However, clinicians were unsure how much effort they were expected to make:

Clinician: *However, I do find it quite difficult when you have this list of phone numbers and you don't get a hold of someone, you know? And, like, how long do you keep on trying?*

Clinician: *You don't even know if the number's still current.*

Clinician: *Yes.*

Clinician: *That's right.*

Clinician: *Yes. That's where we need a policy.*

Second, some clinicians suggested it was hard to be enthusiastic about PN for chlamydia because it was not as serious a disease as gonorrhea, syphilis, or human immunodeficiency virus infection—"they're kind of high-end STIs." With these diseases clinicians described extensive PN:

> *If you had something, you know, that was more serious, for want of a better word ... we would then go all out to really make sure that person contacted their contacts ... but I think with chlamydia, because it's a bit familiar and common ... if the young person says yeah, she'll contact so-and-so, we just leave it at that a lot of the time.*

Finally, clinicians in focus groups expressed a low level of confidence in their ability to tell if their PN efforts were successful. They rarely knew if clients had notified partners, and if they had, if those partners had sought treatment:

> *Yes, because we don't know whether those contacts are going to their GP and getting tested or treated and unless we – unless everyone brings that contact in to us, we don't really know the outcome.*

Although some clinicians described formally following up with returning clients, most had to rely on trusting that clients were disclosing contacts to the clinician and that they would inform their contacts. Some did not feel confident that this trust was well placed:

> *Although sometimes I, kind of, get a sense that, "Oh, yes, you know, I'll tell everybody," and you just think, I don't really know if you are going to do that.*

Partner Notification Facilitators

In focus groups, clinicians identified two key PN facilitators. First, many clinicians said that when they first discuss screening, they also prepare clients for the possibility of having to inform a contact of a positive result:

> *And making sure they're aware of that, that we have to contact trace, before they have the test. That can be useful if they're already prepared for that, if they get a positive test.*

Clinicians described various strategies for explaining the importance of PN, including telling clients it was their legal or moral responsibility. Explaining the logic of PN was especially important where a client's relationship with the sexual contact may be acrimonious:

> *And if you do not treat him, then he might give to other women, other women may give to other men and other men may end up give you - give back to you in the end.*

The second facilitator was the low level of stigma associated with chlamydia among younger clients. This meant they were less embarrassed about having to inform a sexual contact and contacts were likely receptive to the information:

> *This reflects on the education that's out there for young people. When I did talk about contact tracing to one girl, she just gave me the mobile numbers of all the men she'd had sex with in the last six months ... when I did phone each of these people, none of them were surprised or shocked or disbelieving about it. They said, "Yes, okay. Well, I'll go to my doctor." I was very surprised at how accepting the contacts were about this communication from me.*

Providing easy access to screening for partners, clinics also let clients and partners know how easy and noninvasive specimen collection was:

> *Rather than thinking they're going to have something stuck in their urethra or - I say to them, even the GP might not even need to actually even examine your partner if he doesn't have symptoms. He might just have to really just wee in the jar and send it off for a test. It's that easy."*

Directing the Client to the Let Them Know Website

Of the participating clinicians, 23.4% (36/154) reported they always direct their clients to the LTK website, 26.0% (40/154) usually, 37.7% (58/154) sometimes, and 13.0% (20/154) never. According to profession, 49% of both doctors and nurses reported they always or usually directed their clients to LTK website. Always or usually directing the client to the LTK website varied by state or territory: 77.5% (31/40), 66.7% (2/3), 57.1% (8/14), 53.9% (7/13), 47.6% (10/21), 35.0% (7/20), 32.4% (11/34), and 0% (0/9). Also within states, where there were a number of family planning clinics, there was variation. In the state or territory where 0% always or usually directed clients to the website, two thirds of clinicians (66.7%) reported never directing clients to the website and others reported sometimes.

It was clear that while some clinicians promoted the use of the LTK website when they delivered training, awareness among family planning clinicians was not consistent. Some had never heard of it, others were unsure how it worked:

> *I need to go and look up, Let Them Know ...'Cause I haven't seen or heard of those before. So I should go and do that ...Leave that with us and we'll have a look."*

Integration of the Let Them Know Website Into Routine Practice

The focus groups revealed the ways in which the LTK website was integrated into routine practice. For example, it could be raised during the phone call when a patient was being informed by the clinician of their positive chlamydia result:

> I mean I usually tell them about that they need to notify their partners, and just check if they're happy to do it or not...but they're usually like, "Who knows where they are?" So I usually say there is a website, "Do you want it now or do you want to talk when you come in for your treatment?"

Clinicians also described introducing the LTK website during a face-to-face consultation, when it could be demonstrated to the client or used on the spot:

Clinician: *I often get them to do the Let Them Know site right there at the time because it's easy to just pull it up while they're screened.*

Clinician: *Yeah. While they're there. Yeah.*

Clinician: *By the time you've pulled it up and said, "Well, this is how you do it and there it is." It's like, "come on, give me the number."*

Clinician: *I say, "Have you got his number?" And they've got their phone there always.*

Clinician: *Let's do it.*

Clinician: *They've got the number in their phone. You just do it.*

In focus groups, clinicians who did use the LTK website reported it was especially useful for patients who found it difficult to contact partners (these patients may previously have asked clinicians to make direct contact on their behalf):

> I think that service [LTK] was good because I remember a couple of years back I think on two or three occasions I phoned a partner at the client's request because they had a name and a phone number but didn't want to do it themselves. And in those cases, now, they'll use the Let Them Know website. But I agree, the majority of people say, "No, I'm going to tell them. I want to talk to them," you know?

Among survey respondents who never directed their client to the LTK website, 90% said they would like access to this resource.

Let Them Know Letters and Fact Sheets

Clinicians liked LTK for the letters and fact sheets;

> I think that strategy of the letters that are accessible on the web are very useful for GPs to know about, to give the positive patient to give to their partner. And I think those. You know the letters that say - that they give to their boyfriend, then the boyfriend can take it to the GP. He doesn't even have to say anything. And the letter says, "This patient's partner has been diagnosed with chlamydia." And it actually tells the

GP what to treat with them as well. It says, "We suggest that you treat with, and test, and then based on the results, do further contact tracing." And there's an information sheet for the patients as well about what chlamydia is. But I like that actual letter for the GP.

Discussion

To our knowledge this is the first study to evaluate PN practices and the use of the LTK website by family planning clinicians in Australia. We demonstrated that most clinicians take responsibility for supporting their clients to inform their sexual partners and the LTK website was widely used to achieve this goal. Almost half of the clinicians always or usually referred clients to the website, but with considerable variation across clinics in Australian states and territories. The LTK website was considered a useful tool, particularly when clients do not feel comfortable talking to partners about chlamydia.

We found that although family planning clinicians believed PN was primarily the client's responsibility, nearly all supported clients to inform their partners, thus understanding barriers to PN is important. Our survey showed that more than half of the family planning clinicians reported they found PN generally difficult as clients do not always feel comfortable talking to partners about chlamydia and often do not like to name their partners. Other barriers included uncertainty about what was expected of clinicians in relation to PN and doubt about the importance of PN for chlamydia. Facilitators of PN included preparing the client for a positive result when tested, letting clients and partners know how easy specimen collection was, and low levels of stigma about chlamydia in the community. The finding that clinicians were uncertain about what was expected is consistent with surveys of general practitioners in Australia where 45% (105/232) of clinicians were unsure how best to assist their patients with PN with considerable variation in the way PN was undertaken [10].

Despite these barriers, the LTK website was widely used for PN among family planning clinicians, and focus groups revealed that the LTK website was especially useful for patients who found it difficult to contact partners, patients who may previously have relied on clinicians making contact on their behalf. However, there was variation in LTK website use across different states and territories, which may reflect in part the place where it was developed and its subsequent adaption. The website was adapted first in Victoria and New South Wales, with clinics in both states having high proportions of clinicians who reported using LTK always or usually. However, there were clinics in other states where the website was adapted for use in later years, which also had higher proportion of clinicians using LTK, suggesting that the place of development may have played a role in greater use among clinicians, but that other factors also contributed. The LTK website was not formally promoted across Australia; rather, clinics that were aware of it integrated it into clinical practice. Also, the focus groups suggested variation in the formality of PN guidelines and PN training updates.

Despite clinicians recommending clients to use the LTK website, it may not necessarily translate to clients using the website or their partners seeking assessment and treatment. It was raised as a key barrier to PN generally, in that clinicians were unable to judge how successful the activity was. For example, evaluation of the "suggest a test" website in 2 cities in the Netherlands demonstrated that, of those intending to use the website, 23% notified partners using suggest a test and 20% of partners notified by suggest a test subsequently consulted a sexual health clinic [14]. To overcome this, some clinicians mentioned they often asked the client to use the LTK website during the follow-up treatment consultation, providing reassurance to the clinicians that PN had occurred. These findings have implications for general practice, where clinicians have reported not knowing how best to support patients with PN and time would generally be more limited than family planning clinics [10].

The uptake of the PN website among family planning clinicians is far greater than in Australian general practice. A recent study showed only 26% of Australian general practitioners always or usually directed clients diagnosed with chlamydia to the LTK website, compared with 49% in this study [25]. This discrepancy may reflect family planning clinicians' greater expertise in sexual health and participation in meetings and conferences where evaluations of the LTK website and other similar resources were presented. Considering most chlamydia diagnoses occur in general practice [26,27], more proactive approaches may be needed to raise awareness about the LTK website among general practitioners. Sexual health clinics are also a setting where many STI diagnoses occur, but to our knowledge there is no published information on LTK website use among clinicians in this setting. Recently an Internet-based resource and PN service for STIs called "Better to know" was also developed specifically for Aboriginal people including gender-specific information, with the option to anonymously notify partners by SMS text messaging or email [28]. Thus any promotion of such Internet-based services among clinicians should consider the options (LTK, WhyTest, Better to know) available for different target groups.

Strengths and Limitations

The survey had a number of strengths. First, because of questionnaire administration via the clinic representative and reminder, the response rate was much higher than other postal chlamydia knowledge and practice surveys in Australian general practice settings [29-31]. Second, focus groups provided a deeper understanding of how the LTK website was integrated into routine practice and why it wasn't. A number of limitations should also be noted. First, family planning clinicians are only a subset of Australian sexual health providers, and results may not be generalizable to other clinical settings. Second, the clinicians who did not respond to surveys or participate in focus groups may have included more part-time clinicians who had different knowledge, attitudes, and practices than those who did participate. Also, as we did not explore quantitatively whether clinicians knew about the LTK website in the survey (only if they used it), we could not formally assess if there was an association between awareness of the website and referring to it. Third, we did not collect information about state- or territory-based legislation on ability of nurses to supply treatment for people diagnosed with chlamydia. Although this could be a reason for some difference in the use of LTK between states, we do not think it would be a major one. Finally, as the study was conducted in the context of a broader study of chlamydia testing and management practices at Australian family planning clinics, only a subset of questions and a fraction of focus group interview time were dedicated to PN and the LTK website.

Conclusions

In conclusion, the study has demonstrated that the LTK website has become an important PN tool for family planning clinicians, that it has the potential to become part of routine practice. To raise awareness of the LTK website among clinicians, the tool should be specifically mentioned in all clinic protocols and other clinical resources [29,30] and regular organizational newsletters. Also, at the bottom of the pathology reports with positive chlamydia test result, there could be a link added to the Australian STI guidelines that include information about the LTK website. The study also highlights the need for further training and education about PN generally to highlight the importance of PN for chlamydia and information (or a brief algorithm) on how best clinicians could assist their clients with PN. Further research is needed to determine the efficacy of the tool in regard to treatment of the partner and reinfection of the index case.

Conflicts of Interest

None declared.

References

1. Centers for Disease Control and Prevention. Atlanta: U.S. Department of Health and Human Services; 2015. Sexually transmitted disease surveillance 2014 URL: http://www.cdc.gov/std/stats14/ [accessed 2015-12-16] [WebCite Cache ID 6do1w2mga]

2. European Centre for Disease Prevention and Control. Stockholm: ECDC; 2015. Annual epidemiological report 2014 - Sexually transmitted infections, including HIV and blood-borne viruses URL: http://ecdc.europa.eu/en/press/news/_layouts/forms/News_DispForm.aspx?List=8db7286c-fe2d-476c-9133-18ff4cb1b568&ID=1173 [accessed 2015-12-16] [WebCite Cache ID 6do2BHf71]

3. The Kirby Institute. Sydney: UNSW Australia; 2014. HIV, viral hepatitis and sexually transmissible infections in Australia annual surveillance report 2014 URL: https://kirby.unsw.edu.au/surveillance/2014-annual-surveillance-report-hiv-viral-hepatitis-stis [accessed 2015-12-16] [WebCite Cache ID 6do2J5k29]

4. Yeung AH, Temple-Smith M, Fairley CK, Vaisey AM, Guy R, Law MG, et al. Chlamydia prevalence in young attenders of rural and regional primary care services in Australia: a cross-sectional survey. Med J Aust 2014 Feb 17;200(3):170-175. [doi: 10.5694/mja13.10729]

5. Ferreira A, Young T, Mathews C, Zunza M, Low N. Strategies for partner notification for sexually transmitted infections, including HIV. Cochrane Database Syst Rev 2013;10:CD002843. [doi: 10.1002/14651858.CD002843.pub2] [Medline: 24092529]

6. Althaus CL, Heijne JCM, Herzog SA, Roellin A, Low N. Individual and population level effects of partner notification for Chlamydia trachomatis. PLoS One 2012 Dec;7(12):e51438 [FREE Full text] [doi: 10.1371/journal.pone.0051438] [Medline: 23251534]

7. RACGP Clinical guidelines. East Melbourne: Royal Australian College of General Practitioners; 2012. Guidelines for preventive activities in general practice, 8th edition URL: http://www.racgp.org.au/your-practice/guidelines/redbook/ [accessed 2015-12-16] [WebCite Cache ID 6do2QGEhZ]

8. Nwokolo NC, Dragovic B, Patel S, Tong CY, Barker G, Radcliffe K. 2015 UK national guideline for the management of infection with Chlamydia trachomatis. Int J STD AIDS 2016 Mar;27(4):251-267. [doi: 10.1177/0956462415615443] [Medline: 26538553]

9. Workowski KA, Bolan GA. MMWR Recommendations and Reports. 2015. Sexually transmitted diseases treatment guidelines, 2015 URL: http://www.cdc.gov/mmwr/preview/mmwrhtml/rr6403a1.htm [accessed 2016-06-14] [WebCite Cache ID 6iG8NbaYw]

10. Bilardi JE, Hopkins CA, Fairley CK, Hocking JS, Tomnay JE, Pavlin NL, et al. Innovative resources could help improve partner notification for chlamydia in primary care. Sex Transm Dis 2009 Dec;36(12):779-783. [doi: 10.1097/OLQ.0b013e3181b357f6] [Medline: 19704390]

11. Bilardi JE, Fairley CK, Hopkins CA, Hocking JS, Sze JK, Chen MY. Let Them Know: evaluation of an online partner notification service for chlamydia that offers E-mail and SMS messaging. Sex Transm Dis 2010 Sep;37(9):563-565. [doi: 10.1097/OLQ.0b013e3181d707f1] [Medline: 20351621]

12. Huffam S, Fairley CK, Chung M, Sze JK, Bilardi J, Chen MY. Facilitating partner notification through an online messaging service: Let Them Know. Sex Health 2013;10(4):377. [doi: 10.1071/sh13007]

13. Bourne C, Zablotska I, Williamson A, Calmette Y, Guy R. Promotion and uptake of a new online partner notification and retesting reminder service for gay men. Sex Health 2012 Sep;9(4):360-367. [doi: 10.1071/SH11132] [Medline: 22877596]

14. Götz HM, van Rooijen MS, Vriens P, Op de Coul E, Hamers M, Heijman T, et al. Initial evaluation of use of an online partner notification tool for STI, called 'suggest a test': a cross sectional pilot study. Sex Transm Infect 2014 May;90(3):195-200. [doi: 10.1136/sextrans-2013-051254] [Medline: 24391062]

15. Kerani RP, Fleming M, DeYoung B, Golden MR. A randomized, controlled trial of inSPOT and patient-delivered partner therapy for gonorrhea and chlamydial infection among men who have sex with men. Sex Transm Dis 2011 Oct;38(10):941-946. [doi: 10.1097/OLQ.0b013e318223fcbc] [Medline: 21934569]

16. Levine D, Woodruff AJ, Mocello AR, Lebrija J, Klausner JD. inSPOT: the first online STD partner notification system using electronic postcards. PLoS Med 2008 Oct 21;5(10):e213 [FREE Full text] [doi: 10.1371/journal.pmed.0050213] [Medline: 18942887]

17. Plant A, Rotblatt H, Montoya J, Rudy E, Kerndt P. Evaluation of inSPOTLA.org: an internet partner notification service. Sex Transm Dis 2012;39(5):341-345. [doi: 10.1097/OLQ.0b013e31820ef796] [Medline: 22504593]

18. Rietmeijer CA, Westergaard B, Mickiewicz TA, Richardson D, Ling S, Sapp T, et al. Evaluation of an online partner notification program. Sex Transm Dis 2011 May;38(5):359-364. [doi: 10.1097/OLQ.0b013e31820ef796] [Medline: 21343844]

19. Falasinnu T, Gustafson P, Gilbert M, Shoveller J. Risk prediction in sexual health contexts: protocol. JMIR Res Protoc 2013 Dec;2(2):e57 [FREE Full text] [doi: 10.2196/resprot.2971] [Medline: 24300284]

20. Ten HG, Hoebe CJ, van Bergen JE, Brouwers EE, Ruiter RA, Kok G. The influence of two different invitation letters on Chlamydia testing participation: randomized controlled trial. J Med Internet Res 2014 Jan;16(1):e24 [FREE Full text] [doi: 10.2196/jmir.2907] [Medline: 24480721]

21. Bowring AL, Goller JL, Gouillou M, Harvey C, Bateson D, McNamee K, et al. Chlamydia testing and retesting patterns at family planning clinics in Australia. Sex Health 2013 Mar;10(1):74-81. [doi: 10.1071/SH11187] [Medline: 23256944]

XSL•FO

RenderX

22.   Yin R. Qualitative research from start to finish. New York: Guilford Press; 2010.
23.   Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
24.   Australasian Society for HIV Medicine. Sydney: Australasian Society for HIV Medicine; 2010. Australasian contact tracing manual. A practical handbook for health care providers managing people with HIV. viral hepatitis, STIs and HIV-related tuberculosis URL: http://ctm.ashm.org.au/ [accessed 2015-12-16] [WebCite Cache ID 6do1zAzI1]
25.   Lorch R, Yeung A, Hocking J, Wood A, Vaisey A, Fairley C. Are Australian general practitioners and practice nurses equipped for increased chlamydia testing? Findings from ACCEPT. 2012 Presented at: International Union against Sexually Transmitted Infections World Congress; October 15-17, 2012; Melbourne.
26.   Andersen B, Eidner PO, Hagensen D, Lomborg S, Hoff G. Opportunistic screening of young men for urogenital Chlamydia trachomatis infection in general practice. Scand J Infect Dis 2005;37(1):35-39. [doi: 10.1080/00365540510026418] [Medline: 15764188]
27.   Fairley CK, Hocking J, Gunn J, Chen MY. No barriers to chlamydia testing in sexually active young women. Med J Aust 2005 Nov 21;183(10):548-549. [Medline: 16296981]
28.   Better to know. URL: http://www.bettertoknow.org.au/ [accessed 2015-12-16] [WebCite Cache ID 6do5luIo2]
29.   Khan A, Hussain R, Plummer D, Minichiello V. Willingness to offer chlamydia testing in general practice in New South Wales. Aust N Z J Public Health 2006 Jun;30(3):226-230. [doi: 10.1111/j.1467-842X.2006.tb00862.x]
30.   Heal C, Cheffins T, Larkins S, Buhrer-Skinner M, Spillman M. Genital Chlamydia trachomatis infection - a study of general practice management in northern Queensland. Aust Fam Physician 2012 Jul;41(7):519-521 [FREE Full text] [Medline: 22762073]
31.   Temple-Smith MJ, Mak D, Watson J, Bastian L, Smith A, Pitts M. Conversant or clueless? Chlamydia-related knowledge and practice of general practitioners in Western Australia. BMC Fam Pract 2008;9:17 [FREE Full text] [doi: 10.1186/1471-2296-9-17] [Medline: 18307819]

Abbreviations

LTK: Let Them Know
PN: partner notification
SMS: short message service
SRH: sexual and reproductive health
STI: sexually transmitted infection

XSL•FO
RenderX

Original Paper

Exposure to Internet-Based Tobacco Advertising and Branding: Results From Population Surveys of Australian Youth 2010-2013

Sally Dunlop[1,2], B Psych (Hons), PhD (Psych); Becky Freeman[3,4], BSc, MSc, PhD; Donna Perez[1], BSc, MComm

[1]Division of Cancer Screening and Prevention, Cancer Institute New South Wales, Sydney, Australia

[2]Sydney School of Public Health, University of Sydney, Sydney, Australia

[3]Prevention Research Collaboration, Sydney School of Public Health, University of Sydney, Sydney, Australia

[4]Charles Perkins Centre, University of Sydney, Sydney, Australia

Corresponding Author:
Sally Dunlop, B Psych (Hons), PhD (Psych)
Division of Cancer Screening and Prevention
Cancer Institute New South Wales
Level 9, 8 Central Ave,
Australian Technology Park
Sydney, 2015
Australia
Phone: 61 (02) 8374 3665
Fax: 61 (02) 8374 3600
Email: sally.dunlop@cancerinstitute.org.au

Abstract

Background: Since legislation prohibiting tobacco advertising in traditional media, online communication platforms and social media have become one of the few avenues for the tobacco industry to promote its products to Australians. Little is currently known about the exposure of young people to these new media promotions.

Objective: To measure exposure to Internet-based tobacco advertising and branding among Australian youth, identify common formats of branding encountered, and examine the association between exposure and smoking susceptibility.

Methods: The Tobacco Promotion Impact Study is a repeat cross-sectional telephone survey of young people (12-24 years) in 2 Australian states, conducted yearly from 2010 to 2013 (total n=8820). The survey included questions about past-month exposure to Internet-based tobacco advertising and tobacco company branding. Changes in levels of exposure, characteristics of exposed youth, and the association between exposure and smoking susceptibility were explored.

Results: Past-month exposure to Internet-based tobacco advertising and branding among young people increased over the years of the survey (advertising: 21% in 2010 to 29% in 2013; branding: 20% in 2010 to 26% in 2013). The participants who were younger, female, from lower socioeconomic status, and never-smokers were more likely to report exposure. Facebook was the most commonly cited platform for encountering tobacco branding in 2013 (22% of all branding). Compared with young people interviewed in 2013, participants in 2010 were significantly less likely to report exposure to tobacco branding on social media (odds ratio [OR] 0.26, 95% CI 0.20-0.33, $P<.001$) or 2011 (OR 0.46, 95% CI 0.37-0.57, $P<.001$). Among never-smokers aged 12-17 years, exposure to online advertising and branding (OR 1.32, 95% CI 1.11-1.57, $P=.002$) or branding alone (OR 1.39, 95% CI 1.10-1.77, $P=.007$) were significant predictors of smoking susceptibility.

Conclusions: Ensuring tobacco advertising bans are inclusive of Internet-based media is essential. Given the global nature of Internet-based content, cooperation among signatory nations to the World Health Organization Framework Convention Alliance on Tobacco Control will be necessary.

KEYWORDS

youth; tobacco; social media; advertising

XSL•FO
RenderX

Introduction

It is well established that the implementation of a comprehensive tobacco advertising ban is a crucial element of effective tobacco control [1]. Legislation that only limits certain types of tobacco advertising, promotion, and sponsorship enables the tobacco industry to shift resources to unregulated forms of marketing [2]. Equally, the growth in the number and accessibility of new media channels creates opportunities for the tobacco industry to promote its branding and products. The accelerated uptake of communications technology, particularly the use of online social media platforms such as Facebook, has fueled calls for increased surveillance of online tobacco advertising and improved knowledge of the impact it may be having on consumers [3].

The emerging body of research examining prosmoking imagery and advertising online is primarily descriptive in nature [4]. Case study [5,6] and content analysis research [7-9] of new media have shown a proliferation of prosmoking messages, imagery, and tobacco brand promotion. Although direct promotion of tobacco via advertisements is not permitted by the owners of many social media sites (eg, Facebook [10]), the potential for tobacco companies to use these sites to raise the visibility of their products and promote tobacco use remains [11-13]. Tobacco companies can still operate branded pages and channels on social media portals, either directly or through advertising firms that can include product updates, images, videos, and links to real-life events. A recent analysis of 70 cigarette brands on social media revealed more than 120,000 video clips on YouTube and 238 Facebook fan pages with more than 1 million "likes," indicating high user interaction [14].

Given the explosive growth in social media use coupled with the ubiquitous uptake of Internet-enabled mobile phones [15,16], determining levels of exposure to this type of tobacco promotion is essential if tobacco advertising bans are to keep pace with modern marketing methods. Although previous research has revealed that both young people [17] and adults [18] are regularly exposed to Internet-based tobacco advertising, very little published research has determined the level of exposure occurring on popular social media websites. Data collected from US school students in 2011 showed that 11% of youth had received advertisements or promotions from tobacco companies via Facebook or Myspace [19]; however, this does not capture the more indirect forms of promotion that are common on social media.

The potential impact of exposure to online tobacco advertising and promotion on youth is of particular concern. US data show that 92% of teenagers aged 13-17 years are online daily, with 24% reporting they are online almost "constantly." Survey data

from 2015 showed that Facebook is the most used social media site among American teenagers, ages 13 to 17 years, with 71% using the site [20]. Globally, Facebook is the world's most popular social media site, with 1.55 billion monthly active users as of September 2015 [21]. Although a large body of evidence has demonstrated the link between exposure to traditional tobacco branding and smoking susceptibility [22], much remains to be learned about the effects of exposure to tobacco branding in the digital space. One study has shown that exposure to tobacco advertising via Facebook or Myspace was associated with protobacco beliefs and willingness to try smoking among young never-smokers [19], but no study to date has examined the effect of exposure to online tobacco company branding in general, which is the more common form of tobacco promotion on social media.

Australia is known as a "dark market" for tobacco products, with increasingly prohibitive advertising restrictions since the 1970s. Bans on television and radio advertisements were followed by bans on outdoor advertising and sponsorship of sporting events in the 1980s, advertising in the print media and retail point-of-sale in the 1990s, and point-of-sale tobacco displays from 2010. In 2013, Australia introduced the world-first plain packaging legislation in which all forms of branding were removed from tobacco packs in an effort to reduce one of the last forms of tobacco promotion in Australia. Although the global nature of the Internet makes regulating online advertising challenging, amendments to the Tobacco Advertising Prohibition Act made it an offense for any person to publish tobacco advertising on the Internet or other electronic media from Australia, from September 2012 [23]. These regulations also set out requirements for Internet point-of-sale advertisements: they must be in a plain, text-only format with no product images, inclusive of graphic health warnings and accompanied by warnings about age restrictions on sales. No research to date has investigated Australians' exposure to online tobacco advertising and branding in the context of these evolving restrictions.

The primary objective of this study was to assess the exposure of Australian youth to online tobacco advertising and promotion and determine whether exposure has changed in recent years in relation to changes in opportunities for tobacco promotion (outlined in Table 1). We sought to measure exposure to online tobacco advertisements, as well as to more general tobacco company branding, and we profiled youth most likely to be exposed. Additionally, we tracked any changes in the locations where branding was encountered, including social media sites. Finally, we aimed to determine if the established association between exposure to tobacco marketing and smoking susceptibility among youth [24] was also evident with exposure to online tobacco advertising and tobacco branding.

Table 1. Timing of restrictions on tobacco advertising and promotions in relation to Tobacco Promotion Impact Study.

Year	Month	TPIS[a] waves	Restrictions
2010	January		Point-of-sale display ban – NSW[b] large retailers
	June	Wave 1	
	July		Point-of-sale display ban – NSW small retailers
2011	June	Wave 2	
	November		Point-of-sale display ban – QLD[c] all retailers
2012	June	Wave 3	
	September		Internet advertising ban
	October		Plain packaging introduced
2013	June	Wave 4	

[a]TPIS: Tobacco Promotion Impact Study.

[b]NSW: New South Wales.

[c]QLD: Queensland.

Methods

Data for this study come from the Tobacco Promotion Impact Study (TPIS), conducted in the Australian states of New South Wales (NSW) and Queensland (QLD). The study has a repeat cross-sectional design with yearly telephone surveys conducted in June of each year from 2010 to 2013 (total n=8820). The TPIS monitors adolescents' and young adults' (12-24 years) exposure to tobacco promotions in a range of places, as well as smoking-related cognitions and behaviors. Households were recruited using random digit dialing and participants within households were recruited using random selection (selecting the *n*th oldest eligible person aged 12 to 24 years). From 2010 to 2012, recruitment was conducted using landline phone numbers only. In 2013, because of concerns about the increasing proportion of Australian homes without a landline phone number (from 17% in 2010 to 22% in 2012) [25], a supplemental sample of participants was also recruited through random-digit dialing to mobile phone numbers. Use of this supplemental sample is described below. Permission was obtained from parents of 12- to 15-year-olds before conducting each interview. Cooperation rates averaged 70% among eligible respondents. When taking into account households of unknown eligibility, response rates averaged 42% (American Association for Public Opinion Research Response Rate #3) [26]. The study was approved by the NSW Population and Health Services Research Ethics Committee.

Measures

Exposure to Internet-Based Tobacco Promotion

To take into account direct advertising as well as the forms of more indirect promotion encountered on social media sites,

exposure to both online tobacco advertising and online tobacco branding was assessed. All respondents were asked, "In the past month, how often have you seen any promotions or advertising for cigarettes or other tobacco products in the following places?". The list of possible places included "the Internet" (online advertising). They were also asked, "In the past month, how often have you seen cigarette brands, tobacco company names, or logos on the internet?" (online branding). Responses to both questions were "never," "rarely," "sometimes," and "often." Responses to both questions were dichotomized because of negative skew (0=never or rarely vs 1=sometimes or often). Responses were also combined to indicate whether an individual was (1) never/rarely exposed to advertising or branding, (2) sometimes/often exposed to both advertising and branding, (3) sometimes/often exposed to advertising only, or (4) sometimes/often exposed to branding only.

In order to explore types of branding encountered, young people exposed to branding were asked in what formats the cigarette brands, tobacco company names, or logos were encountered on the Internet. Responses were recorded verbatim and matched to a list of possible websites or types of online advertising. We also combined responses in order to report on the proportion of respondents seeing (1) branding in advertisements (pop-up advertisements, banner advertisements, Google advertisements, website advertisements); (2) branding on social media (Twitter, Facebook, YouTube, Myspace); (3) branding in personal communications (email, instant messenger, forums); and (4) branding on content-controlled websites (news sites, sports sites, blogs, gaming sites, Yahoo, Ninemsn; see Figure 1).

Figure 1. Summary of questions and derived variables relating to online tobacco promotion exposure. Survey questions are denoted by "Q"; derived variables are numbered.
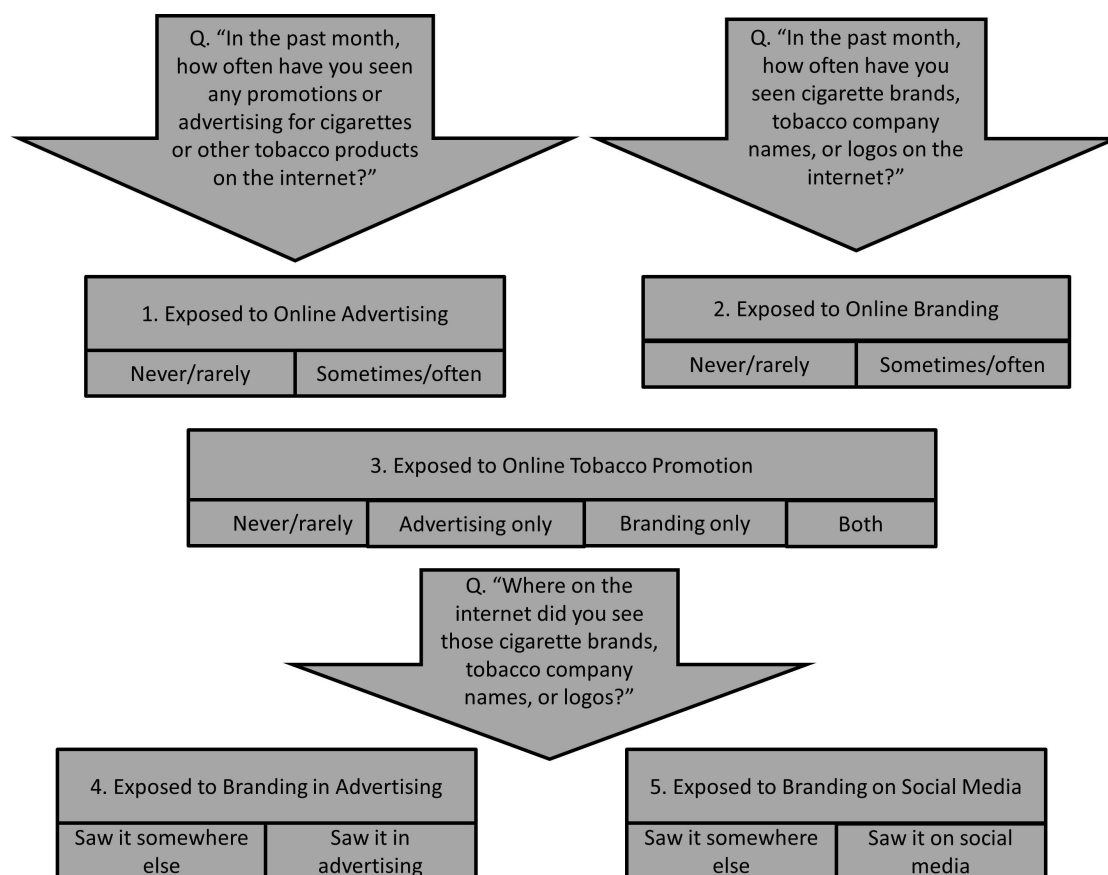
Current Smoking

Respondents were asked if they had ever had a puff of a cigarette, how many cigarettes they had smoked in their lifetime, and if they had smoked in the past month. Based on stage models of smoking uptake [27], they were classified as follows: (1) never-smokers (never taken a puff); (2) experimenters (smoked less than 5 cigarettes ever, or smoked 5-100 cigarettes in their lifetime but not in the past month); (3) current smokers (smoked more than 5 cigarettes in their lifetime, and smoked in the past month); or (4) ex-smokers (smoked more than 100 cigarettes in their lifetime but not in the past month).

Smoking Susceptibility

Never-smokers were asked a series of validated questions to determine their susceptibility to smoking in the future [28,29]. Participants were classified as nonsusceptible if they answered "definitely no" to each of the following questions: "Do you think that you will try cigarettes sometime soon?"; "Do you think you will smoke a cigarette sometime in the next year?"; and "If a friend offered you a cigarette, would you try it?" (response options: 1=definitely no, 2=probably no, 3=probably yes, 4=definitely yes). Participants who did not answer each of those questions with "definitely no" were classified as susceptible.

Smoking Exposure

Respondents reported on the number of current smokers in their household and how many of their five closest friends smoked.

Average Daily Internet Use

Respondents were asked, "How much time do you spend on average per day on the Internet, if at all?" Responses were recorded in minutes and divided by 60 to represent hours per day.

Demographics

Age, sex, state of residence, and year of interview were included. Postcodes were used with the Socio-Economic Indexes for Areas (SEIFA) [30] to indicate low (quintiles 4-5) or moderate-high (quintiles 1-3) socioeconomic status (SES).

Statistical Analysis

We first conducted logistic regression analyses to explore changes over time in exposure to (1) online tobacco advertising and (2) online tobacco branding. Each logistic regression model included year of interview, demographics (age, sex, SES, state), Internet use, smoking exposures (friends, household), and smoking status as predictors. Because of the low number of ex-smokers in the sample, and the similar demographic profile of ex-smokers and current smokers, these groups were combined for these analyses. These models also identified individual characteristics of exposed youth.

Next, we examined changes over time in the format that tobacco branding was encountered online. The overall number of young people who reported seeing each of the branding formats was relatively small (ranging from 1 to 829), so we report only on those mentioned by at least 5% of the sample who recalled

seeing branding. We used Pearson chi-square tests to detect significant differences in exposure to specific branding formats, as well as the types of branding, over survey years. The overall numbers of young people who reported seeing branding in personal communications (n=90) or on content-controlled websites (n=226) were small, therefore they were not investigated further. Logistic regression analyses were conducted to predict encountering branding through (1) advertisements and (2) social media. Year of interview, demographics, smoking status, smoking exposures, and Internet use were entered as predictors.

Finally, we explored whether smoking susceptibility was associated with exposure to online advertising and tobacco branding. As well as testing whether smoking susceptibility was associated with exposure to online advertising, we were also interested in whether exposure to tobacco company branding in the absence of advertising would be associated with susceptibility. Therefore, we created a 3-level variable classifying participants as having been exposed in the past month (1) never/rarely to online tobacco advertising or branding, (2) sometimes/often to tobacco advertising (with or without branding), or (3) sometimes/often exposed to tobacco company branding but not advertising. This variable was entered as a predictor in a logistic regression model predicting smoking susceptibility, with demographic characteristics, year of interview, smoking exposures, and Internet use as covariates. Because the established link between exposure to protobacco marketing and smoking susceptibility applies primarily to adolescents [24], we conducted this analysis separately for 12- to 17-year-olds and 18- to 24-year-olds.

The supplemental mobile phone sample was added in 2013 to assess whether any changes in outcomes between years of the survey were due to changes in the characteristics of the population covered by landlines. Previous studies have found that adding a mobile component to a landline population survey gives a more representative sample [31], but it also has the potential to result in changes to population estimates that are a consequence of the design change, rather than a real change [32]. Comparing both the landline-only and the dual-frame (landline and mobile) samples with previous years' samples

allows this issue to be explored. Therefore, all analyses in this study were conducted twice. The first set of analyses used the landline sample only, comparing differences between years while minimizing bias due to changes in sampling. The second set of analyses used the dual-frame sample for 2013; comparing differences between years while minimizing the influence of the changing composition of a sample recruited via landline only. The results from the second set of analyses are only reported when the pattern of results differ from the first.

The gender distribution of this sample was relatively consistent with population parameters as defined by Australian Bureau of Statistics data [30]. There were, however, some discrepancies in the age distribution, particularly a slight overrepresentation of 16- to 19-year-olds but underrepresentation of 20- to 24-year-olds. Given these discrepancies, data were weighted to the NSW and QLD populations of 12- to 24-year-olds for age, sex, and region distributions from Census data [30] using poststratification weights. In the set of analyses including the 2013 mobile phone supplement, additional weighting was used to account for telephone status (landline only, mobile phone only, or dual user). All analyses were conducted using Stata v11.1 [33].

Results

Sample characteristics and exposure to online tobacco promotion for each survey wave are listed in Table 2(with landline and dual-frame samples for 2013 shown separately). The samples in each year of the survey were similar in terms of age and sex. There was a significant difference in SES, with the highest proportion of respondents from a moderate-high SES area in 2011. There was a significant difference in smoking status in the landline sample: current smoking decreased from 16% to 12% over the years of the survey. In the dual-frame sample, there was an increase from 12% in 2012 to 16% in 2013. Similarly, the number of smoking friends and household members also decreased significantly in the landline sample but increased in the dual-frame sample in 2013. Average daily Internet use increased significantly, from 2.43 hours in 2010 to 3.28 hours in 2013 (3.33 in the dual-frame sample).

Table 2. Sample characteristics.

Characteristic	2010 (n=2000)	2011 (n=2010)	2012 (n=2003)	2013 landline (n=2001)	P [a]	2013 dual-frame (n=2807)	P [a]
Age in years, N(%) [b]							
12-15	669 (32)	649 (32)	619 (32)	684 (31)		833 (31)	
16-19	826 (31)	833 (30)	855 (30)	799 (31)		1046 (31)	
20+	505 (37)	528 (38)	529 (38)	518 (39)	.974	928 (39)	.950
Sex, N(%) [b]							
Female	975 (49)	990 (49)	992 (49)	980 (49)		1325 (49)	
Male	1025 (51)	1021 (51)	1011 (51)	1021 (51)	>.99	1482 (51)	>.99
State, N(%) [b]							
NSW[c]	1000 (50)	1004 (50)	1000 (50)	1001 (50)		1407 (50)	
QLD[h]	1000 (50)	1000 (50)	1000 (50)	1000 (50)	>.99	1400 (50)	>.99
SES [d,f], N(%) [b]							
Low	557 (28)	497 (25)	578 (29)	536 (26)		735 (28)	
Moderate-high	1443 (72)	1514 (75)	1425 (71)	1465 (74)	.011	2056 (72)	.015
Smoker, N(%) [b]							
Current	293 (16)	243 (13)	220 (12)	207 (12)		369 (16)	
Never	1178 (56)	1278 (61)	1276 (61)	1376 (64)		1769 (60)	
Former	53 (4)	42 (3)	36 (2)	36 (3)		70 (3)	
Experimenter	476 (25)	448 (24)	471 (25)	382 (21)	<.001	599 (22)	<.001
Internet-based tobacco promotion exposure, N(%) [b]							
Never/rarely	1330 (70)	1273 (66)	1200 (62)	1158 (61)		1651 (59)	
Ads[g] and branding[e]	218 (11)	266 (13)	313 (15)	335 (16)		442 (16)	
Ads only[e]	208 (10)	240 (12)	245 (12)	272 (13)		395 (15)	
Branding only[e]	200 (10)	205 (10)	223 (11)	215 (10)	<.001	293 (11)	<.001
Friends who smoke, mean (SD)	1.27 (1.60)	1.10 (1.51)	1.05 (1.47)	0.93 (1.41)	<.001	1.11 (1.52)	<.001
Household members who smoke, mean (SD)	0.52 (0.86)	0.47 (0.86)	0.49 (1.01)	0.42 (0.82)	.048	0.51 (0.99)	<.001
Internet use in hours, mean (SD)	2.43 (2.36)	2.63 (2.49)	2.90 (2.59)	3.28 (2.96)	<.001	3.33 (3.00)	<.001

[a]*P* values from chi-square tests for differences between proportions or analysis of variance tests for differences between means.

[b]Numbers are unweighted, percentages are weighted.

[c]NSW: New South Wales;.

[d]Based on postal code.

[e]Sometimes/rarely exposed.

[f]SES: socioeconomic status

[g]ads: advertisements.

[h]QLD: Queensland.

Exposure to Internet-Based Tobacco Advertising and Branding

There were significant differences in recent exposure to Internet-based tobacco promotion across the years of the survey, with the proportion of the sample never or rarely exposed decreasing from 70% in 2010 to 61% in 2013; 59% in dual-frame sample. In 2013, 16% of participants were recently exposed to both advertising and branding, 13% exposed to advertising only; 15% in dual-frame sample, and 10% to branding only; 11% in dual-frame sample.

Table 3 shows the proportions of youth exposed to Internet-based tobacco advertising and branding, along with the results from the logistic regression analyses predicting exposure in the landline samples. Controlling for demographic and smoking characteristics, youth interviewed in 2010 or 2011 were significantly less likely to have recently been exposed to Internet-based tobacco advertising than those interviewed in 2013. In the landline sample, there was no significant difference in the likelihood of being exposed to advertising or branding between 2012 and 2013. However, in the dual-frame sample, youth interviewed in 2012 were significantly less likely than those interviewed in 2013 to report exposure (odds ratio [OR] 0.85, 95% CI 0.73-0.98, P=.024). In the model predicting exposure to tobacco company branding, youth interviewed in 2010 and 2011 were significantly less likely to report recent exposure than those interviewed in 2013. There was no change in the proportion exposed to branding between 2013 and 2012 in either the landline or dual-frame samples.

Table 3. Proportions of youth (unadjusted) with exposure to Internet-based tobacco advertising and branding, and results from logistic regression analyses predicting exposure (results from landline sample).

Characteristic	Exposed[a] to tobacco advertising/promotion (n=7856)				Exposed[a] to tobacco company branding (n=7858)			
	%[b]	OR[c,d]	95% CI		P	%[b]	OR[c]	95% CI		P
Year										
2010	21	0.66	0.57	0.77	<.001	20	0.71	0.61	0.83	<.001
2011	24	0.81	0.70	0.94	.006	22	0.82	0.70	0.95	.010
2012	27	0.92	0.80	1.07	.284	26	0.98	0.84	1.14	.780
2013	29	(ref[e])				26	(ref)			
Age, years										
12-15	33	2.12	1.81	2.49	<.001	32	2.58	2.18	3.04	<.001
16-19	27	1.61	1.40	1.86	<.001	25	1.79	1.54	2.08	<.001
20+	17	(ref)				15	(ref)			
Sex										
Female	29	(ref)				26	(ref)			
Male	22	0.71	0.64	0.79	<.001	21	0.77	0.69	0.86	<.001
State										
NSW[f]	26	(ref)				23	(ref)			
QLD[g]	25	0.96	0.87	1.07	.499	24	1.05	0.94	1.18	.348
SES [h,j] Low	28	(ref)				27				
Moderate-high	24	0.83	0.74	0.94	.003	22	0.81	0.72	0.92	<.001
Smoking										
Never-smoker	30	(ref)				27	(ref)			
Experimenter	22	0.81	0.70	0.94	.006	21	0.87	0.74	1.01	.065
Current or ex-smoker	12	0.37	0.29	0.47	<.001	13	0.44	0.35	0.57	<.001
Friends who smoke	N/A[i]	1.08	1.03	1.13	.001	N/A	1.10	1.05	1.15	<.001
Household members who smoke	N/A	1.08	1.01	1.15	.030	N/A	1.12	1.05	1.20	.001
Internet, hours	N/A	1.04	1.02	1.06	<.001	N/A	1.04	1.02	1.06	<.001

[a]Exposure=sometimes or often exposed versus never or rarely.

[b]Percentages are weighted.

[c]Odd ratios are from multivariable analyses.

[d]OR: odds ratio.

[e]ref: reference category.

[f]NSW: New South Wales.

[g]QLD: Queensland.

[h]SES: socioeconomic status.

[i]N/A: not applicable.

[j]Based on postal code.

There were many similarities in the characteristics of youth most likely to be exposed to Internet-based advertising and branding: participants who were younger, female, and from lower SES areas were more likely to report exposure. Current smokers were less likely to be exposed than never-smokers. There were positive associations between friends' smoking,

household members' smoking, average daily Internet use, and both types of exposure.

Format of Tobacco Branding

Across all years of the survey, when asked where they had seen tobacco company branding on the Internet, the most common answer among youth was that they did not know (Table 4). However, this proportion decreased significantly from 40% in

2010 to 28% in 2013 (29% dual-frame). In 2013, the most common place to report seeing tobacco branding was on Facebook, followed by pop-up messages, banner advertisements, YouTube, and Google advertisements. Chi-square analyses showed that exposure to branding on Facebook and YouTube increased significantly over the years of the study, while exposure to branding on Google advertisements decreased.

Table 4. Format of branding encountered among youth who reported seeing Internet-based tobacco branding.

Format[a]	2010 (n=850)	2011 (n=967)	2012 (n=999)	2013 landline (n=1033)	P [b]	2013 dual-frame (n=1384)	P [b]
Pop-up messages	19%	21%	20%	19%	.632	20%	.643
Banner ads[c]	16%	17%	19%	17%	.314	16%	.203
Google ads	4%	4%	7%	3%	.004	3%	.002
Facebook	9%	15%	21%	22%	<.001	22%	<.001
YouTube	2%	3%	9%	12%	<.001	11%	<.001
Don't know	40%	32%	27%	28%	<.001	29%	<.001

[a]Only formats with at least 5% of sample naming them are included.

[b]P value from Pearson chi-square tests for proportions.

[c]Ads: advertisements.

The proportions of youth who saw Internet-based branding and reported that they saw it in advertising or in social media, along with the results from the logistic regression analyses predicting these exposures, are listed in Table 5 (landline sample). Controlling for differences in demographic and smoking characteristics, youth interviewed in 2012 were significantly more likely than youth interviewed in 2013 to report encountering branding in advertising. When the model was run with the dual-frame sample, youth interviewed in 2011 (OR 1.21, 95% CI 1.01-1.46, P=.043) and 2012 (OR 1.26, 95% CI 1.05-1.51, P=.014) were significantly more likely than those interviewed in 2013 to report encountering branding in advertising. Conversely, youth interviewed in 2010 or 2011 were significantly less likely to have encountered branding on social media than those interviewed in 2013 (same pattern of results obtained in the landline and dual-frame samples). Males were less likely than females to have encountered branding in advertising or social media. Current and ex-smokers were less likely than never-smokers to have encountered branding in advertising. Participants with more friends who smoke, and those with higher Internet use, were more likely to have encountered tobacco company branding on social media.

Table 5. Proportions of youth (unadjusted) exposed to different formats of Internet-based tobacco branding, and results from logistic regression analyses predicting exposure (results from landline sample).

Characteristic		Exposed to branding in advertising versus exposed elsewhere (n=3849)				Exposed to branding on social media versus exposed elsewhere (n=3849)			
		%[a]	OR[b,c]	95% CI		P	%[a]	OR[b]	95% CI		P
Year											
	2010	37	0.99	0.82	1.21	.959	10	0.26	0.20	0.33	<.001
	2011	40	1.16	0.96	1.40	.119	17	0.46	0.37	0.57	<.001
	2012	42	1.21	1.01	1.45	.043	27	0.83	0.68	1.01	.066
	2013	37	(ref[e])				31	(ref)			
Age, years											
	12-15	43	1.18	0.97	1.44	.105	20	1.36	1.05	1.75	.018
	16-19	38	1.03	0.86	1.24	.724	25	1.41	1.13	1.76	.002
	20+	36	(ref)				20	(ref)			
Sex											
	Female	41	(ref)				23	(ref)			
	Male	37	0.84	0.73	0.96	.010	20	0.83	0.70	0.97	.022
State											
	NSW[f]	39	(ref)				21	(ref)			
	QLD[g]	39	0.99	0.87	1.14	.912	22	1.02	0.86	1.20	.842
SES [d,h]											
	Low	39	(ref)				22	(ref)			
	Moderate-high	39	0.99	0.85	1.15	.885	22	1.03	0.86	1.24	.731
Smoking											
	Never-smoker	41	(ref)				21	(ref)			
	Experimenter	38	0.99	0.83	1.19	.932	21	0.94	0.75	1.18	.613
	Current or ex-smoker	28	0.70	0.51	0.95	.021	30	1.39	0.99	1.96	.060
Friends who smoke		N/A[i]	0.96	0.91	1.02	.178	N/A	1.15	1.07	1.23	<.001
Household members who smoke		N/A	1.02	0.94	1.10	.610	N/A	1.03	0.94	1.12	.578
Internet, hours		N/A	0.99	0.97	1.02	.663	N/A	1.06	1.03	1.09	<.001

[a]Percentages are weighted.

[b]Odds ratios are from multivariable analyses.

[c]OR: odds ratio.

[d]Based on postal code.

[e]ref: reference category.

[f]NSW: New South Wales.

[g]QLD: Queensland.

[h]SES: socioeconomic status.

[i]N/A: not applicable.

Association Between Exposure to Tobacco Advertising or Branding and Smoking Susceptibility

Results from the logistic regression analysis predicting smoking susceptibility among never-smokers are shown separately for adolescents and young adults in Table 6. For adolescents, compared with those never or rarely exposed to online tobacco promotion, those exposed to online tobacco advertising (with or without branding) as well as those exposed to tobacco company branding only were more likely to be susceptible to smoking. These effects were apparent when controlling for the influence of age, household members and friends smoking, year of interview, and average daily Internet use. For young adults, there were no associations between exposure to online tobacco promotions and smoking susceptibility. These results were the same in both the landline and dual-frame samples.

Table 6. Proportions of nonsmoking youth (unadjusted) susceptible to smoking, and results from logistic regression analyses predicting smoking susceptibility (results from landline sample).

Characteristic		12- to 17-year-olds (n=3377)				18- to 24-year-olds (n=1594)			
		%[a]	OR[b,c]	95% CI		*P*	%[a]	OR[b]	95% CI		*P*
Exposure										
	Never/rarely	23	(ref[f])				13	(ref)		
	Exposed[d] to online ads[g] and branding	29	1.32	1.11	1.57	.002	16	1.30	0.92	1.83	.137
	Exposed[d] to online branding only	30	1.39	1.10	1.77	.007	11	0.76	0.44	1.31	.315
Year										
	2010	24	1.00	0.79	1.25	.967	10	0.64	0.41	1.00	.049
	2011	25	1.06	0.85	1.32	.610	13	0.80	0.53	1.22	.301
	2012	29	1.22	0.98	1.52	.069	15	0.99	0.67	1.46	.955
	2013	25	(ref)				15	(ref)		
Sex										
	Female	23	(ref)				12	(ref)		
	Male	28	1.30	1.11	1.53	.001	15	1.43	1.05	1.93	.021
State										
	NSW[c]	25	(ref)				13	(ref)		
	QLD[h]	26	0.99	0.85	1.16	.931	14	1.11	0.82	1.51	.480
SES [e,i]										
	Low	24	(ref)				15	(ref)		
	Moderate-high	26	1.13	0.94	1.36	.187	13	0.88	0.63	1.23	.445
Friends who smoke		N/A[j]	1.23	1.13	1.34	<.001	N/A	1.04	0.93	1.16	.495
Household members who smoke		N/A	1.16	1.05	1.29	.005	N/A	1.22	1.00	1.50	.050
Internet, hours		N/A	1.05	1.01	1.09	.009	N/A	0.99	0.94	1.04	.697

[a]Percentages are weighted.

[b]Odds ratios are from weighted analyses.

[c]OR: odds ratio.

[d]Exposure=sometimes or often exposed.

[e]Based on postal code.

[f]ref: reference category.

[g]ads: advertisements.

[h]QLD: Queensland.

[i]SES: socioeconomic status.

[j]N/A: not applicable.

Discussion

This study is the first to assess levels of exposure to online tobacco promotion in Australia, a notoriously "dark market." The results suggest that not only is tobacco advertising and branding commonly encountered by young Australians, with almost a third of the youth surveyed in 2013 exposed, it is also increasing on social media, specifically Facebook.

Over the years of the study, exposure to online tobacco advertising and tobacco branding increased from 2010 to 2012.

Concurrent changes to Australian tobacco advertising legislation included moving to retail tobacco displays bans and plain packaging of tobacco products. It has been noted that, as opportunities for tobacco promotion in one domain are restricted, the tobacco industry's efforts in other domains increase [2]. In our study, the observed increases in exposure to online promotion were independent of increases in Internet use, or any changes in sample composition. These results may suggest that tobacco company efforts at attracting young Australians are being directed toward Internet-based advertising in the face of increasing restrictions on other forms of promotion, although

XSL•FO
RenderX

this should be verified with monitoring of online advertising. Exposure to online advertising and branding appeared to plateau in 2013, concurrent with the national legislation banning Internet-based tobacco advertising originating from Australia. Nevertheless, and perhaps unsurprisingly given the borderless nature of the Internet, online tobacco promotions remained readily accessible to Australian youth.

This study extended previous research [19,29] by measuring exposure not only to tobacco advertisements, but also to tobacco company branding in general. While there was a degree of overlap between exposure to online tobacco advertising and tobacco branding, around 10% of youth reported being exposed to tobacco company branding in the absence of advertising. Of the participants who reported seeing tobacco branding online in 2013, around one-third reported seeing it in "traditional" Internet advertising formats such as pop-up advertisements, banner advertisements, sponsored search engine results, and website advertisements. This was a significant decrease from 2012, which might indicate a small effect of the national legislation introduced at the end of 2012.

Concurrent with the decrease in exposure to tobacco branding in traditional forms of online advertising, there were increases in exposure to branding on social media sites. Around a third of youth who saw online tobacco branding in 2013 reported seeing it on social media. Australians are prolific Facebook users, with 13.2 million users as of June 2014, making it one of the most popular websites in Australia [34]. It may be somewhat expected then that tobacco branding was most commonly reported as being seen on Facebook. Although Facebook prohibits advertisements that directly promote the sale of tobacco products, it does not prohibit advertisements that promote the use of tobacco products among like-minded individuals. Additionally, advertisements in this context are very narrowly defined, including only paid advertising that is purchased and prepared through the Facebook advertising portal. Any tobacco promotions appearing as unpaid content would be exempt from this policy. This presents a unique challenge for regulators, as unlike more traditional forms of Internet-based advertising, tobacco marketing on social media sites is less amenable to regulation and more difficult to directly attribute to tobacco companies [35]. Innovative approaches are likely to be required in order to determine the origins of this type of content, perhaps by engaging computer science and technology experts who can accurately navigate online networks.

There were also rapid increases in the proportion of youth seeing tobacco branding on YouTube. YouTube is also an exceptionally popular website, with an estimated 1 billion unique visitors from around the world every month [36]. Prosmoking imagery and tobacco promotions have been well documented on YouTube [14,37]. Again, like Facebook, YouTube does not allow tobacco products to be advertised on the site. The definition of advertising on YouTube is incredibly narrow, however, and only applies to *paid* forms of promotion on the site, such as advertisements embedded in popular videos or advertisements that appear for certain key word searches. British American Tobacco (BAT), for example, has its own YouTube channel, WelcomeToBAT, which includes videos outlining BAT's public positions on harm reduction, illicit tobacco,

marketing, and sustainable farming [38]. The broad definition of tobacco marketing outlined in Article 13 of the World Health Organization (WHO) Framework Convention on Tobacco Control (FCTC) could encompass this type of material because it has "the aim, effect or likely effect of promoting a tobacco product or tobacco use either directly or indirectly" [39].

In this study, participants most likely to recall seeing online tobacco advertising or branding were younger (12-15 years old) and/or female. Future research might explore potential reasons for this, including whether the tobacco industry is targeting younger people with media placement strategies, whether the advertising has been designed to appeal most to these demographics, or whether these groups are particularly sensitive to branding that speaks to evolving identities. Of note, nonsmoking youth were more likely to remember seeing tobacco advertising and branding than current smokers. Contrary to tobacco industry claims that any promotions are aimed at creating brand loyalty and switching among current smokers [40], online advertisements are reaching young people with no experience of smoking.

The fact that younger participants and nonsmokers were the most likely to report exposure to online tobacco advertising and branding is particularly concerning, as the younger never-smokers who remembered seeing tobacco advertising, promotions, or branding were more likely to be susceptible to smoking. This relationship was apparent even when controlling for smoking among family and friends. Building on the well-established link between exposure to tobacco company marketing and smoking susceptibility [2], this study is the first to establish a link between smoking susceptibility and exposure to online tobacco advertising as well as online tobacco branding of the type found on social media sites. We did not find an association between tobacco advertising or branding with smoking susceptibility for the older group of nonsmokers, indicating that other factors may be more important influencers of smoking susceptibility at that age.

Strengths and Limitations

The strengths of this study include the collection of data over 4 years and in 2 states, resulting in a large and relatively representative sample. A wide range of covariates was used in all analyses, limiting the likelihood that observed changes in exposure over time were due to sample variations. Additionally, the inclusion of the mobile phone supplementary sample in 2013 allowed us to verify that the patterns of results we observed in the landline sample were primarily apparent in the dual-frame sample, reducing concerns about the use of sampling bias due to landline recruiting for 2010-2012. This study extends previous research on exposure of young people to online tobacco promotion by including exposure to tobacco branding as well as advertising and by identifying specific formats of branding encountered.

Limitations of the study include relying on self-reported exposure to online material—it can be difficult to remember where precisely something was seen online, as evidenced by the high number of participants stating that they did not know where they encountered tobacco branding. Given that much of what we see online is unlikely to be recalled, we are potentially

underestimating rates of actual exposure. The findings of this study should therefore be interpreted alongside existing investigations about the number and nature of advertisements and promotions, particularly on social media sites [14]. Additionally, young people have been known to underreport smoking-related behaviors over the phone compared with when they self-complete a survey [41]—this might have slightly diminished estimated rates of smoking and smoking susceptibility, particularly in the younger age group, but this effect would have been consistent across years. Finally, the observed association between exposure to online tobacco promotion and smoking susceptibility is cross-sectional in nature, and longitudinal data would be needed to investigate the order of effects. However, the inclusion of a large number of appropriate covariates demonstrates that this association exists independently of the influence of the exposure of young people to peer and family smoking.

Conclusions

The relatively common experience of exposure to online tobacco advertising, promotion, and branding among Australian youth reinforces the importance of comprehensive restrictions on Internet tobacco promotion, as well as strong counter-advertising initiatives. The WHO FCTC recognizes that cross-border promotions are a threat to domestic laws that ban tobacco advertising [42]. This is particularly true for online promotions where the borderless nature of the Internet will require cooperation among parties to the WHO FCTC in order for tobacco advertising bans to be effective. Establishing mechanisms where WHO FCTC parties can monitor, report,

and act on promotions that leak across borders is paramount. Even in such a climate of cooperation, it is likely that the greatest challenge in monitoring and enforcing restrictions on Internet tobacco promotions will be to linking such promotions to the tobacco industry, especially on social media. Accordingly, an expert advisory group has been proposed to keep WHO FCTC parties up to date on relevant developments in technology in cross-border tobacco advertising, promotion, and sponsorship, and in best practices for responding to these forms of promotion [43]. There is also a need to continually assess the media strategies used in counter-advertising so that antitobacco messages are reaching young people in the digital spaces where they are likely to be encountering tobacco promotion.

Our results stress the need for continued research and surveillance of tobacco marketing that is penetrating new and underregulated digital media. In order to continue downward trends in smoking and the denormalization of smoking among youth, online advertising and promotions need to be subject to more comprehensive restrictions. There is a misconception that online advertising is a weaker or less penetrative form of marketing [44] and is simply used to augment more traditional offline media promotions. Evidence from the alcohol control field demonstrates that online advertising reduces the effectiveness of regulations banning offline advertising because online advertising replaces, rather than simply complements, offline advertising [45]. Given the demonstrated effect of tobacco promotion on tobacco uptake by young people [24], increased efforts to restrict youth exposure to tobacco promotion through these new media outlets are critical.

Authors' Contributions

SD and DP conceived the study, SD analyzed the data, SD and BF drafted the manuscript, and all authors contributed critical revisions to the manuscript.

Conflicts of Interest

None declared.

References

1. World Health Organisation TFI. MPOWER. 2014. Enforce bans on tobacco advertising, promotion and sponsorship URL: http://www.who.int/tobacco/mpower/enforce/en/ [accessed 2016-02-03] [WebCite Cache ID 6f0aiCJ1c]

2. National Cancer Institute. The Role of the Media in Promoting and Reducing Tobacco Use. Bethesda, MD: National Cancer Institute, National Institutes of Health; 2008.

3. Ribisl KM, Jo C. Tobacco control is losing ground in the Web 2.0 era: invited commentary. Tob Control 2012 Mar;21(2):145-146. [doi: 10.1136/tobaccocontrol-2011-050360] [Medline: 22345237]

4. Forsyth SR, Kennedy C, Malone RE. The effect of the internet on teen and young adult tobacco use: a literature review. J Pediatr Health Care 2013;27(5):367-376. [doi: 10.1016/j.pedhc.2012.02.008] [Medline: 22521497]

5. Freeman B, Chapman S. British American Tobacco on Facebook: undermining Article 13 of the global World Health Organization Framework Convention on Tobacco Control. Tob Control 2010 Jun;19(3):e1-e9 [FREE Full text] [doi: 10.1136/tc.2009.032847] [Medline: 20395406]

6. Seidenberg AB, Rees VW, Connolly GN. Swedish Match marketing on YouTube. Tob Control 2010 Dec;19(6):512-513. [doi: 10.1136/tc.2010.038919] [Medline: 20930059]

7. Bromberg JE, Augustson EM, Backinger CL. Portrayal of smokeless tobacco in YouTube videos. Nicotine Tob Res 2012 Apr;14(4):455-462 [FREE Full text] [doi: 10.1093/ntr/ntr235] [Medline: 22080585]

8. Freeman B, Chapman S. Is "YouTube" telling or selling you something? Tobacco content on the YouTube video-sharing website. Tobacco control 2007;16:207-210.

9. BinDhim NF, Freeman B, Trevena L. Pro-smoking apps for smartphones: the latest vehicle for the tobacco industry? Tob Control 2014 Jan;23(1):e4. [doi: 10.1136/tobaccocontrol-2012-050598] [Medline: 23091161]

10. Facebook. Facebook Advertising Policies. 2016. URL: https://www.facebook.com/unsupportedbrowser [accessed 2016-04-27] [WebCite Cache ID 6h63ocI5c]

11. Elkin L, Thomson G, Wilson N. Connecting world youth with tobacco brands: YouTube and the internet policy vacuum on Web 2.0. Tob Control 2010 Oct;19(5):361-366. [doi: 10.1136/tc.2010.035949] [Medline: 20739706]

12. Carroll MV, Shensa A, Primack BA. A comparison of cigarette- and hookah-related videos on YouTube. Tob Control 2013 Sep;22(5):319-323 [FREE Full text] [doi: 10.1136/tobaccocontrol-2011-050253] [Medline: 22363069]

13. Freeman B, Chapman S. British American Tobacco on Facebook: undermining Article 13 of the global World Health Organization Framework Convention on Tobacco Control. Tob Control 2010 Jun;19(3):e1-e9 [FREE Full text] [doi: 10.1136/tc.2009.032847] [Medline: 20395406]

14. Liang Y, Zheng X, Zeng DD, Zhou X, Leischow SJ, Chung W. Exploring how the tobacco industry presents and promotes itself in social media. J Med Internet Res 2015;17(1):e24 [FREE Full text] [doi: 10.2196/jmir.3665] [Medline: 25608524]

15. Pew Research Center. Global Digital Communication: Texting, Social Networking Popular Worldwide 2012. URL: http://www.pewglobal.org/2011/12/20/global-digital-communication-texting-social-networking-popular-worldwide/ [accessed 2016-02-03] [WebCite Cache ID 6f0bMwx91]

16. Pew Research Center. Social Networking Popular Across Globe. URL: http://www.pewglobal.org/2012/12/12/social-networking-popular-across-globe/ [accessed 2016-02-03] [WebCite Cache ID 6f0c6Tdif]

17. Agaku IT, King BA, Dube SR. Trends in exposure to pro-tobacco advertisements over the Internet, in newspapers/magazines, and at retail stores among U.S. middle and high school students, 2000-2012. Prev Med 2014 Jan;58:45-52. [doi: 10.1016/j.ypmed.2013.10.012] [Medline: 24183778]

18. US Surgeon General, Centers for Disease Control & Prevention. Adult Awareness of Tobacco Advertising, Promotion, and Sponsorship - 14 Countries. JAMA 2012;308(6):561-562.

19. Cavazos-Rehg PA, Krauss MJ, Spitznagel EL, Grucza RA, Bierut LJ. Hazards of new media: youth's exposure to tobacco Ads/promotions. Nicotine Tob Res 2014 Apr;16(4):437-444 [FREE Full text] [doi: 10.1093/ntr/ntt168] [Medline: 24163285]

20. Pew Research Center. Teen, Social Media and Technology Overview 2015 URL: http://www.pewinternet.org/2015/04/09/teens-social-media-technology-2015/ [accessed 2016-02-02] [WebCite Cache ID 6f0cNHyBI]

21. Facebook. Statistics 2015. URL: http://newsroom.fb.com/company-info/ [accessed 2016-02-02] [WebCite Cache ID 6f0cW6cd1]

22. DiFranza JR, Wellman RJ, Sargent JD, Weitzman M, Hipple BJ, Winickoff JP, Tobacco Consortium, Center for Child Health Research of the American Academy of Pediatrics. Tobacco promotion and the initiation of tobacco use: assessing the evidence for causality. Pediatrics 2006 Jun;117(6):e1237-e1248. [doi: 10.1542/peds.2005-1817] [Medline: 16740823]

23. Australian Government. Tobacco Advertising Prohibition Amendment Act 2012. URL: https://www.legislation.gov.au/Details/C2012A00005 [accessed 2016-04-27] [WebCite Cache ID 6h64U4VyZ]

24. Lovato C, Watts A, Stead LF. Impact of tobacco advertising and promotion on increasing adolescent smoking behaviours. Cochrane Database Syst Rev 2011(10):CD003439. [doi: 10.1002/14651858.CD003439.pub2] [Medline: 21975739]

25. Australian Communications Media Authority (ACMA). Communications Report 2010-11 series. 2011. Report 2 - Converging communications channels: Preferences and behaviours of Australian communications users URL: http://www.acma.gov.au/webwr/_assets/main/lib410148/report2-convergent_comms.pdf [accessed 2016-04-27] [WebCite Cache ID 6h64kHBvu]

26. American Association for Public Opinion Research. Standard definitions: final dispositions of case codes and outcome rates for surveys. Lenexa, Kansas: AAPOR; 2008.

27. Mayhew KP, Flay BR, Mott JA. Stages in the development of adolescent smoking. Drug Alcohol Depend 2000 May 1;59 Suppl 1:S61-S81. [Medline: 10773438]

28. Pierce JP, Choi WS, Gilpin EA, Farkas AJ, Merritt RK. Validation of susceptibility as a predictor of which adolescents take up smoking in the United States. Health Psychol 1996 Sep;15(5):355-361. [Medline: 8891714]

29. Dube SR, Arrazola RA, Lee J, Engstrom M, Malarcher A. Pro-tobacco influences and susceptibility to smoking cigarettes among middle and high school students--United States, 2011. J Adolesc Health 2013 May;52(5 Suppl):S45-S51 [FREE Full text] [doi: 10.1016/j.jadohealth.2012.07.007] [Medline: 23601611]

30. Australian Bureau of Statistics. Socio-economic indexes for areas (SEIFA). In: 2006 Census of Population and Housing. Canberra, Australia: Commonwealth of Australia; 2008.

31. Barr ML, van Ritten Jason J, Steel DG, Thackway SV. Inclusion of mobile phone numbers into an ongoing population health survey in New South Wales, Australia: design, methods, call outcomes, costs and sample representativeness. BMC Med Res Methodol 2012;12:177 [FREE Full text] [doi: 10.1186/1471-2288-12-177] [Medline: 23173849]

32. Barr ML, Ferguson RA, Steel DG. Inclusion of mobile telephone numbers into an ongoing population health survey in New South Wales, Australia, using an overlapping dual-frame design: impact on the time series. BMC Res Notes 2014;7:517 [FREE Full text] [doi: 10.1186/1756-0500-7-517] [Medline: 25113743]

33. StataCorp. 2009. Stata Statistical Software V11 URL: http://www.stata.com/ [accessed 2016-05-01] [WebCite Cache ID 6hAoqzy8y]

34. Social Media Statistics Australia - June 2014. URL: http://www.socialmedianews.com.au/social-media-statistics-australia-june-2014/ [accessed 2016-02-03] [WebCite Cache ID 6f0eQ9Wdg]

35. Freeman B. New media and tobacco control. Tob Control 2012 Mar;21(2):139-144. [doi: 10.1136/tobaccocontrol-2011-050193] [Medline: 22345236]

36. YouTube. YouTube Statistics. 2015. URL: http://www.youtube.com/yt/press/statistics.html [accessed 2016-02-03] [WebCite Cache ID 6f0h0GWiq]

37. Kim K, Paek H, Lynn J. A content analysis of smoking fetish videos on YouTube: regulatory implications for tobacco control. Health Commun 2010 Mar;25(2):97-106. [doi: 10.1080/10410230903544415] [Medline: 20390676]

38. British American Tobacco. Welcome to BAT URL: https://www.youtube.com/user/WelcomeToBAT [accessed 2016-02-03] [WebCite Cache ID 6f0hG8E5u]

39. World Health Organizatio (WHO). WHO Framework Convention on Tobacco Control, Article 13. 2003. URL: http://www.who.int/tobacco/control/measures_art_13/en/ [accessed 2016-04-27] [WebCite Cache ID 6h65HAK0L]

40. Cummings K, Morley C, Horan J, Steger C, Leavell N. Marketing to America's youth: evidence from corporate documents. Tob Control 2002;11:5-17.

41. Moskowitz JM. Assessment of Cigarette Smoking and Smoking Susceptibility among Youth: Telephone Computer-Assisted Self-Interviews versus Computer-Assisted Telephone Interviews. Public Opinion Quarterly 2004 Dec 01;68(4):565-587. [doi: 10.1093/poq/nfh040]

42. Conference of the Parties to the WHO Framework Convention on Tobacco Control. Measures that would contribute to the elimination of cross-border advertising, promotion and sponsorship (decision FCTC/COP3(14)). 2010. URL: http://apps.who.int/gb/fctc/PDF/cop4/FCTC_COP4_10-en.pdf [accessed 2016-02-03] [WebCite Cache ID 6f0hOlmru]

43. Kenyon A, Liberman J. University of Melbourne Legal Studies Research Paper No 161. Controlling Cross-Border Tobacco: Advertising, Promotion and Sponsorship - Implementing the FCTC (August 2006) URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=927551 [accessed 2016-02-03] [WebCite Cache ID 6f0hkSMx9]

44. Shaw M. Online advertising doesn't work - it's official. Or is it?. URL: http://creativepool.com/magazine/news/advertising/online-advertising-doesnt-work--its-official--or-is-it.3178 [accessed 2016-02-03] [WebCite Cache ID 6f0hvlgvY]

45. Goldfarb A, Tucker C. Advertising Bans and the Substitutability of Online and Offline Advertising. Journal of Marketing Research 2011 Apr;48(2):207-227. [doi: 10.1509/jmkr.48.2.207]

Abbreviations

BAT: British American Tobacco
FCTC: Framework Convention on Tobacco Control
NSW: New South Wales
OR: odds ratio
SES: socioeconomic status
TPIS: Tobacco Promotion Impact Study
WHO: World Health Organization

Original Paper

Health-Specific Information and Communication Technology Use and Its Relationship to Obesity in High-Poverty, Urban Communities: Analysis of a Population-Based Biosocial Survey

Anjali Gopalan[1,2,3], MD, MS; Jennifer A Makelarski[4], PhD, MPH; Lori B Garibay[4], MA, MPH; Veronica Escamilla[4], PhD; Raina M Merchant[3], MD, MS; Marcus B Wolfe Sr[5,6], MSW; Rebecca Holbrook[7], BA; Stacy Tessler Lindau[8,9,10,11], MD, MAPP

[1]Corporal Michael J. Crescenz VA Medical Center, Philadelphia, PA, United States

[2]Kaiser Permanente Northern California, Division of Research, Oakland, CA, United States

[3]Robert Wood Johnson Clinical Scholars Program, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, United States

[4]The University of Chicago, Department of Obstetrics and Gynecology, Chicago, IL, United States

[5]University of Chicago, Chicago, IL, United States

[6]Concordia University-Chicago, Chicago, IL, United States

[7]Holbrook Consulting, Chicago, IL, United States

[8]The University of Chicago, Department of Obstetrics and Gynecology, and Medicine-Geriatrics, Chicago, IL, United States

[9]The University of Chicago Comprehensive Cancer Center, Chicago, IL, United States

[10]The University of Chicago Maclean Center on Clinical Medical Ethics, Chicago, IL, United States

[11]The University of Chicago Urban Health Initiative, Chicago, IL, United States

Corresponding Author:
Anjali Gopalan, MD, MS
Kaiser Permanente Northern California
Division of Research
2000 Broadway
Oakland, CA, 94612
United States
Phone: 1 510 891 3458
Fax: 1 510 891 3508
Email: Anjali.Gopalan@kp.org

Abstract

Background: More than 35% of American adults are obese. For African American and Hispanic adults, as well as individuals residing in poorer or more racially segregated urban neighborhoods, the likelihood of obesity is even higher. Information and communication technologies (ICTs) may substitute for or complement community-based resources for weight management. However, little is currently known about health-specific ICT use among urban-dwelling people with obesity.

Objective: We describe health-specific ICT use and its relationship to measured obesity among adults in high-poverty urban communities.

Methods: Using data collected between November 2012 and July 2013 from a population-based probability sample of urban-dwelling African American and Hispanic adults residing on the South Side of Chicago, we described patterns of ICT use in relation to measured obesity defined by a body mass index (BMI) of ≥ 30 kg/m^2. Among those with BMI≥ 30 kg/m^2, we also assessed the association between health-specific ICT use and diagnosed versus undiagnosed obesity as well as differences in health-specific ICT use by self-reported comorbidities, including diabetes and hypertension.

Results: The survey response rate was 44.6% (267 completed surveys/598.4 eligible or likely eligible individuals); 53.2% were African American and 34.6% were Hispanic. More than 35% of the population reported an annual income of less than US $25,000. The population prevalence of measured obesity was 50.2%. People with measured obesity (BMI≥ 30 kg/m^2) were more likely to report both general (81.5% vs 67.0%, P=.04) and health-specific (61.1% vs 41.2%, P=.01) ICT use. In contrast, among those with measured obesity, being told of this diagnosis by a physician was not associated with increased health-specific ICT use.

XSL•FO
RenderX

People with measured obesity alone had higher rates of health-specific use than those with comorbid hypertension and/or diabetes diagnoses (77.1% vs 60.7% vs 47.4%, P=.04).

Conclusions: In conclusion, ICT-based health resources may be particularly useful for people in high-poverty urban communities with isolated measured obesity, a population that is at high risk for poor health outcomes.

Introduction

Obesity and obesity-related chronic diseases are leading drivers of health care costs in the United States [1]. Over the past 35 years, the prevalence of obesity has more than doubled; currently, 35% of American adults are obese (defined as body mass index, BMI, $\geq$30 kg/m^2) [1]. Certain populations are disproportionately affected by obesity, including African American and Hispanic adults and people living in resource-poor, high-poverty, and more racially segregated urban communities [1-5]. Many major technology corporations, payers, and health care systems are investing in information and communication technologies (ICTs), such as mobile apps and Web-based patient portals, to prevent and better manage obesity and related chronic conditions [6-10]. Growing, but limited, evidence demonstrates that ICT-based interventions can positively affect health behaviors and obesity-related outcomes [6,11,12]. Small, clinic-based trials of mobile apps and other Web-based decision support and monitoring tools have demonstrated improved outcomes for specific chronic health conditions, including short-term weight loss [6,11]. Although these findings are promising, the ability of health-specific ICT-based resources to impact health outcomes will depend not only on efficacy in clinical trials but on actual use among people with obesity, especially those residing in communities with limited health resources [13].

The 2012 Pew Internet Health Tracking Survey examined the relationship between types of ICT use, including seeking information online about conditions, medications, or the experiences of others, and self-reported chronic disease [14]. In this study, controlling for age, income, education, ethnicity, and overall health status, people who self-reported a diagnosis of chronic disease, including hypertension, diabetes, heart, and lung disease, were less likely to report any ICT use (the frequency of these activities was not described) [14]. However, among people who did report ICT use, those with one or more chronic diseases were more likely to report the use of ICT for health-specific reasons compared with those without a chronic condition [14]. Obesity, designated by the American Medical Association as a chronic condition after the study's completion, was not included among the chronic conditions examined in the Pew survey [15]. Also missing from the Pew survey are any biometric data regarding chronic disease status, specifically BMI. A population-based survey that collected individual-level data on general and health-specific ICT use and chronic disease status, including both self-reported obesity diagnoses and objective obesity status (anthropometric measures) presented an opportunity to address these gaps in the Pew data [16].

On Chicago's South Side, 55% of the population (approximately 528,000) lives at or below 200% federal poverty level; 77% of residents are African American, 13% are Hispanic [17]. African American and Hispanic people have disproportionately high rates of obesity [2]. Vital statistics data for the region suggest higher rates of premature and overall mortality related to obesity-related chronic diseases compared with more affluent areas of Chicago [18]. As part of a larger strategy to mitigate health inequities in the region, community leaders, residents, and university researchers with the South Side Health and Vitality Studies conducted a population health survey to establish prevalence estimates for obesity and other chronic diseases and to ascertain what kinds of community-based and ICT resources residents use to manage their health [19]. We hypothesized that residents use ICT resources to substitute for gaps in community-based resources and that ICT use varies by health condition. Understanding the feasibility of ICT-based health resources to reach people with obesity in high-poverty communities is critical to maximizing the potential of these resources to impact health.

In this study, we first describe overall patterns of current ICT use by measured obesity status. Based on the Pew findings for other chronic conditions, we hypothesized that although overall ICT use would be lower among respondents with measured obesity, health-specific ICT use would be higher. Second, among people with measured obesity (BMI$\geq$30 kg/m^2), we described health-specific ICT use comparing those with and without a physician's diagnosis of obesity and by the presence of self-reported comorbid conditions, specifically diabetes and hypertension. The Health Belief Model suggests that willingness to perform a health behavior, such as using health-specific ICT resources, depends on the perceived need for action [20]. Thus, obese people who were never told by a physician they had obesity or an obesity-related comorbid condition may be less motivated to use ICT-based health resources than people with a physician's diagnosis or who have one or more of these comorbid diagnoses [20]. We, therefore, hypothesized that health-specific ICT use would be higher among people with BMI$\geq$30 kg/m^2 who have been diagnosed as obese by a physician and those with measured obesity who also self-report a diagnosis of comorbid hypertension and/or diabetes. Lastly, based on the study findings and extant literature, we propose a conceptual framework describing the relationship between obesity and health-specific ICT use that we hope will inform the design and translation of ICT-based interventions targeting obesity management.

Methods

This analysis is based on data from the South Side Health and Vitality Studies (SSHVS) [16]. SSHVS is a family of interrelated, community-engaged research studies that aim to inform efforts to promote and maintain population health on Chicago's South Side [21]. SSHVS aims to describe population health in the region and the ways in which residents use ICT to access health-related community resources or substitute for gaps in local resources. All participants provided written documentation of informed consent. This study was designed in partnership with volunteer community members [19] and was approved by the University of Chicago Institutional Review Board (IRB). The primary data collection for this research and the activity of the University of Chicago researchers were conducted under the approved University of Chicago IRB protocol. Secondary data analysis was conducted by AG following a human subjects research exemption granted by University of Pennsylvania IRB. Other co-authors had no access to the study subjects or individual-level data.

Study Population

Individuals eligible for this study included those 35 years of age or older, English or Spanish speaking, and residing within the target region.

Sampling

Study participants were sampled from 2 distinct regions, a total of 7 census tracts, on the South Side of Chicago. The northwest region was almost entirely African American (98%), based on 2010 US Census data [22]. The southeast region was majority Hispanic (83%). We employed a single-frame, two-stage sampling design. First, an address-based probability sample of household units was generated by randomly selecting household units from a list of all residential postal addresses in the regions purchased from Marketing Systems Group's Genesys Division, 2012 [23]. Then, if more than one individual in a household was 35 years or older, one individual was randomly selected.

Data Collection

Eligible participants were recruited between November 2012 and July 2013 through mailed letters, telephone calls, and home visits. Once informed consent was obtained, participants took part in an in-person, interviewer-administered structured interview, lasting approximately 1 hour. Participants could choose to complete the interview in English or Spanish. The interview collected sociodemographic characteristics, details on ICT device ownership and use, and self-reported medical history. Physical measures were obtained, including height, weight, waist circumference, systolic and diastolic blood pressure, and finger-stick dried blood specimens.

Defining Information and Communication Technology Use

General ICT use was defined as any use of cell phones for texting, emailing, going online, or downloading apps, or any use of the Internet (accessed via computer or cell phone). Health-specific ICT activities included the following activities: looking up health information online, using a health-related mobile app, Web-based purchasing of medications, Web-based communication with providers, participation in online health support groups, and Web-based management of health records and/or benefits. Questions included on the survey instrument relating to health-specific ICT use were primarily based on a prior national-level survey, the Health Information National Trends Survey, which was modified by the study team to increase cultural appropriateness and aid comprehension (Textbox 1) [24]. On the basis of review of our survey instrument by community informants, an additional question was added asking those who reported looking up health information online the following question: "Do you ever use the Internet to find health information because you did not want to ask a doctor?" This question was included to explore a hypothesis that people might use the Internet for information to avoid embarrassing discussions or to compensate for limited time with health care providers. For all of these health-specific activities, use was primarily defined as any amount of engagement (ranging from daily to less than monthly) or no engagement. The distribution of activity frequency was also described. The dichotomous categorization of health-specific ICT use was done to allow examination of associations between selected activities and participant characteristics (small cell counts would prevent statistical testing) and because the "right" amount of each of the examined activities is unknown and likely varies greatly. This categorization also reflected that used in the aforementioned Pew survey that studied ICT use and self-reported chronic conditions [14].

Textbox 1. Survey instrument used to assess selected health-specific information and communication technology activities (questions 1-5 were modified from the 2003 Health Information National Trends Survey, item HC-26, and question 5a was constructed based on input from community stakeholders).

Instructions: I'm going to list some ways people use the Internet. Some people have done these things and some have not. Please use Card #X to tell me how often you did these things in the past 12 months. In the past 12 months, how often have you...

☐ Every day ☐ At least once a week ☐ At least once a month ☐ <Once a month ☐ Never

1. Bought some kind of medicine online. This includes prescription medicines, over-the-counter medicines, or herbal supplements?

2. Taken part in an online support group for people with a health or medical issue?

3. Used e-mail or the Internet to talk with a doctor or a doctor's office or hospital?

4. Looked for health or medical information online?

5. Looked at or managed your health records online

5a: If yes, do you ever use the Internet to find health information because you did not want to ask a doctor?

☐ Yes ☐ No ☐ Don't Know ☐ Refused

6. Looked at or managed your health benefits, like filing an insurance claim online

---

Defining Obesity and Other Chronic Disease Variables

Body mass index was calculated using measured height and weight collected during in-home interviews [25]. For the primary analysis, measured obesity was defined as a binary variable: nonobese (BMI<30 kg/m^2) and obese (BMI≥30 kg/m^2; Figure 1). Those with missing BMI data (n=12) were excluded. To examine the effect of being diagnosed as obese by a physician, respondents who were categorized as obese by their measured BMI were then stratified by their response to the survey question "Has a medical doctor ever told you that you have excess weight or obesity?" (Figure 1). The resultant binary variable categories were labeled "diagnosed obese" and "undiagnosed obese." To assess differences based on a diagnosis of comorbid chronic conditions, diabetes and hypertension were defined as binary variables based on survey responses to the following two questions: (1) "Has a medical doctor ever told you that you have diabetes?" and (2) "Has a medical doctor ever told you that you have high blood pressure or hypertension?" Given the significant overlap between the 3 examined conditions, the 3 categories were defined as follows: measured obesity only, measured obesity plus self-reported hypertension or diabetes, and measured obesity, self-reported hypertension, and self-reported diabetes.

Figure 1. Population-based probability sample enrollment flowchart. BMI: body mass index.

Statistical Analysis

All analyses for this population-based probability sample were weighted to account for differential selection probabilities and differential nonresponse. The response rate and the cooperation rate were calculated using the American Association for Public Opinion Research (AAPOR) definitions for response rate (RR3) and cooperation rate (COOP3) [26]. The response rate describes the number of completed surveys in relation to the total number of eligible individuals (we assumed an eligibility rate of 63% for 261 individuals of unknown eligibility; Figure 1), whereas the cooperation rate describes the number of completed surveys in relation to the number of eligible individuals ever contacted [26]. Descriptive statistics were calculated to summarize sociodemographic characteristics by obesity status and diagnosis of obesity. Differences in general and health-specific ICT use among the people with and without measured obesity were examined. Among those with measured obesity, the associations between health-specific ICT use and diagnosis of obesity were also compared. Lastly, differences in health-specific ICT use among those with measured obesity only were compared with those reporting a diagnosis of hypertension or diabetes and those with all 3 conditions. Chi-square tests were used to test for statistical differences between or among groups. All data analyses were performed using Stata version 12.1 (2011, StataCorp, College Station, TX, USA).

Results

The response rate was 44.6% (267 completed surveys/598.4 eligible or likely eligible individuals) and the cooperation rate was 61.5% (267 completed surveys/434 eligible individuals contacted). In total, 267 individuals participated in the biosocial study.

Sociodemographic Characteristics and ICT Use by Measured Obesity Status

Associations between sociodemographic characteristics and measured obesity (BMI$\geq$30 kg/m^2) status in the population are summarized in Table 1. Measured obesity was more prevalent among women than men (63.2% vs 36.8%, P=.01). Measured obesity status did not differ by income, education, or race or ethnicity. The majority of people reported seeing a physician in the past year (79.8%) and this did not differ by measured obesity status (82.5% obese vs 78.6% nonobese, P=.55). However, individuals with measured obesity were more likely

to report a source of regular care than those without obesity (96.7% obese vs 88.6% nonobese, P=.04).

The prevalence of general ICT use in the population was high (75.0%) but was more common among people with obesity when compared with people without obesity (81.5% obese vs 67% nonobese, P=.04; Table 2). Half of all respondents (51.7%) reported some type of health-specific ICT use, but it was more common among people with measured obesity (61.1% obese vs 41.2% nonobese, P=.01). The most common health-specific ICT activity in the population was looking up health-related information online (47.2%); this activity was more common among the people with obesity (54.4% obese vs 38.7% nonobese, P=.04). Other health-related activities, including Web-based medical record access (8.8%), participation in online health-related support groups (9.3%), and Web-based communication with providers (9.7%), were infrequent and did not differ by obesity status (Table 2). Health-specific mobile app use was also uncommon (7.6%) and did not differ by measured obesity status. A supplementary table describes the distribution of frequencies for the examined health-specific ICT activities (see Multimedia Appendix 1).

Sociodemographic Characteristics and ICT Use by Obesity Diagnosis

Among people with measured obesity, a physician's diagnosis of obesity was associated with educational attainment; people with obesity who had been diagnosed by a physician were more likely to have achieved a high school diploma or have passed a general educational development (GED) test (41.2% diagnosed vs 24.6% undiagnosed, P=.01) or some college experience (41.2% diagnosed vs 26.0% undiagnosed, P=.01; Table 3). Obese individuals with an obesity diagnosis were also more likely to report a source of regular health care (100% diagnosed vs 89.1% undiagnosed, P=.01). Obesity diagnosis was similar across the other examined sociodemographic characteristics (Table 3). Rates of ICT use by obesity diagnosis were not significantly different for general (85.6% diagnosed vs 72.0% undiagnosed, P=.15) or health-specific use (65.9% diagnosed vs 49.8% undiagnosed, P=.15). Although no statistically significant differences were noted in use of Web-based resources to look up health-related information (60.7% diagnosed vs 39.6% undiagnosed, P=.06), people with diagnosed obesity were more likely to look up information online as a means to avoid asking a doctor (32.0% diagnosed vs 5.0% undiagnosed, P<.001). No other differences in rates of specific health-related ICT activities by obesity diagnosis status were noted.

Table 1. Sociodemographic characteristics of the population and by measured obesity status (results of weighted analysis).

Characteristic	Total population % (95% CI)	Nonobese (BMI[a]<30 kg/m^2) % (95% CI)	Measured obese (BMI≥30 kg/m^2) % (95% CI)	P [b]
	N=267[c]	n=115 44.9 (37.7-52.2)	n=140 50.2 (43.0-57.5)	
Age, years				
35-40	11.8 (7.5-16.0)	9.5 (3.9-15.1)	13.9 (7.3-20.6)	.42
41-50	36.1 (28.8-43.5)	38.1 (26.8-49.3)	35.7 (25.4-45.9)	
51-60	22.4 (17.3-27.6)	20.8 (13.3-28.3)	23.5 (16.1-30.8)	
61-70	13 (8.2-17.7)	17.5 (9.4-25.6)	9.1 (3.3-14.9)	
71+	16.7 (11.1-22.4)	14.1 (6.4-21.8)	17.8 (9.4-26.2)	
Gender				
Male	42.2 (34.9-49.4)	49.5 (38.5-60.5)	36.8 (26.6-47.0)	.01
Female	57.8 (50.6-65.1)	50.5 (39.5-61.5)	63.2 (53.0-73.4)	
Race/ethnicity				
Black, non-Hispanic	53.2 (48.9-57.5)	51.4 (42.4-60.4)	55.2 (46.4-63.9)	.30
Hispanic	34.6 (28.7-40.5)	39.1 (29.0-49.3)	29.4 (19.4-39.4)	
Other	12.2 (7.8-16.6)	9.5 (3.6-15.3)	15.4 (8.4-22.4)	
Income, US $				
<$25K	36.2 (29.5-43.0)	40.7 (30.1-51.3)	34.2 (24.8-43.5)	.59
$25K-$49K	28.4 (22.0-34.7)	24.1 (15.4-32.8)	33.9 (24.4-43.5)	
$50K-$99K	16.7 (11.2-22.2)	14.2 (6.7-21.8)	17.1 (8.9-25.3)	
≥$100K	6.7 (2.8-10.7)	8.4 (1.3-15.6)	5.8 (1.2-10.4)	
Don't know/refused	12.0 (6.6-17.5)	12.5 (3.7-21.4)	9.0 (2.4-15.5)	
Education				
Middle school/ some high school	30.4 (23.5-37.4)	33.3 (22.6-44.0)	27.2 (17.8-36.6)	.55
High school graduate/GED[d]	34.8 (27.8-41.8)	37 (26.7-47.4)	36.2 (26.1-46.3)	
Associates/ some college	34.7 (28.1-41.3)	29.7 (19.9-39.5)	36.6 (27.4-45.9)	
Employment status				
Unemployed	14.3 (9.5-19.2)	15.5 (7.7-23.2)	14.7 (8.0-21.4)	.85
Employed	45.5 (38.2-52.7)	46.4 (35.5-57.3)	42.5 (32.3-52.6)	
Retired	18.7 (13.2-24.1)	18.4 (10.6-26.2)	18.6 (10.6-26.6)	
Unable to work	10.0 (5.4-14.5)	7.1 (0.4-13.8)	12.5 (5.9-19.1)	
Other	11.6 (6.5-16.6)	12.6 (4.3-20.8)	11.8 (4.9-18.7)	
Health insurance				
Uninsured	24.6 (18.1-31.0)	33.8 (23.0-44.5)	17.1 (9.7-24.5)	.01
Medicaid only	8.9 (5.0-12.8)	4.5 (0.8-8.1)	13.7 (6.9-20.6)	
Medicare only	12.4 (8.0-16.8)	15.4 (7.8-23.1)	10.8 (5.7-16.0)	
Private/other	38.0 (31.0-44.9)	35.4 (25.0-45.7)	40.5 (30.5-50.5)	
Multiple	16.2 (10.9-21.4)	11.0 (4.8-17.2)	17.9 (9.8-25.9)	
Physician visit in past year (% yes)	79.8 (73.6-86.0)	78.6 (69.5-87.7)	82.5 (73.6-91.4)	.55

XSL•FO
RenderX

Characteristic	Total population % (95% CI)	Nonobese (BMI[a]<30 kg/m^2) % (95% CI)	Measured obese (BMI≥30 kg/m^2) % (95% CI)	P [b]
	N=267[c]	n=115 44.9 (37.7-52.2)	n=140 50.2 (43.0-57.5)	
Source of regular care (% yes)	91.4 (87.4-95.4)	88.6 (82.2-94.9)	96.7 (93.1-1.0)	.04

[a] BMI: body mass index.

[b]P value for comparison between nonobese and obese populations; significant at level P<0.05.

[c] A total of 12 people were missing a BMI value. These individuals were excluded from the chi-square analysis.

[d] GED: general educational development.

Table 2. Information and communication technology–based activities of the population and by measured obesity status (results of weighted analysis).

ICT[a] activities (% reported yes)	Total population % (95% CI)	Measured nonobese % (95% CI)	Measured obese % (95% CI)	P [b]
	N=267[c]	n=115 44.9 (37.7-52.2)	n=140 50.2 (43.0-57.5)	
General ICT use	75.0 (68.5-81.3)	67.0 (56.7-77.3)	81.5 (73.1-90.0)	.04
Any health-specific use	51.7 (44.5-59.0)	41.2 (30.6-51.9)	61.1 (51.0-71.1)	.01
Seek health info online	47.2 (40.0-54.5)	38.7 (28.1-49.3)	54.4 (44.1-64.6)	.04
Use Web-based resources to avoid asking doctor	20.2 (14.8-25.7)	16.0 (8.9-23.5)	23.9 (15.7-32.1)	.17
Web-based access of health benefit info	12.2 (7.8-16.6)	5.8 (0.9-10.7)	18.6 (11.2-26.1)	.01
Participate in online health support group	9.3 (5.1-13.4)	8.4 (2.8-14.0)	10.5 (4.0-17.0)	.63
Web-based access of health records	8.8 (4.9-12.7)	4.9 (−0.1 to 9.9)	12.7 (6.5-18.9)	.08
Web-based medication purchasing	7.7 (4.3-11.1)	5.1 (0.9-9.3)	10.0 (4.5-15.5)	.17
Web-based communication with providers	9.7 (5.9-13.5)	7.3 (2.5-12.1)	12.4 (6.2-18.6)	.20
Use health-related mobile app	7.6 (3.8-11.4)	5.0 (−0.03 to 10.0)	9.0 (3.6-14.3)	.31

[a] ICT: information and communication technology.

[b]P value for comparison between measured nonobese and obese populations; significant at level P<0.05.

[c] A total of 12 people were missing a body mass index value. These individuals were excluded from the chi-square analysis.

Table 3. Sociodemographic characteristics of the measured obese population and by obesity diagnosis status (results of weighted analysis).

Characteristic	Total measured obese[a] (BMI[b]≥30 kg/m^2) % (95% CI)	Undiagnosed obese % (95% CI)	Diagnosed obese % (95% CI)	P [c]
	n=140	n=43	n=97	
	50.2 (43.0-57.5)	30.1 (20.7-39.5)	69.9 (60.5-79.3)	
Age, years				
35-40	13.9 (7.3-20.6)	12.3 (1.4-23.3)	14.6 (6.3-23.0)	.78
41-50	35.7 (25.4-45.9)	32.2 (14.2-50.3)	37.2 (24.7-49.7)	
51-60	23.5 (16.1-30.8)	26.9 (12.4-41.4)	22.0 (13.5-30.5)	
61-70	9.1 (3.3-14.9)	5.6 (−0.9 to 12.2)	10.5 (2.8-18.3)	
71+	17.8 (9.4-26.2)	22.9 (5.1-40.7)	15.6 (6.6-24.7)	
Gender				
Male	36.8 (26.6-47.0)	45.5 (26.9-64.1)	33.0 (20.7-45.4)	.15
Female	63.2 (53.0-73.4)	54.5 (35.9-73.1)	67.0 (54.6-79.3)	
Race/ethnicity				
Black, non-Hispanic	55.2 (46.4-63.9)	46.3 (30.3-62.3)	59.0 (49.0-69.0)	.32
Hispanic	29.4 (19.4-39.4)	40.5 (22.6-58.4)	24.6 (12.7-36.5)	
Other	15.4 (8.4-22.4)	13.1 (1.4-24.9)	16.4 (7.9-25.0)	
Income, US $				
<$25K	34.2 (24.8-43.5)	32.6 (16.7-48.5)	34.9 (23.3-46.5)	.41
$25K-$49K	33.9 (24.4-43.5)	36.4 (18.0-54.7)	32.9 (21.7-44.0)	
$50K-$99K	17.1 (8.9-25.3)	10.3 (−1.1 to 21.7)	20.1 (9.5-30.6)	
≥$100K	5.8 (1.2-10.4)	4.3 (−1.8 to 10.3)	6.5 (0.4-12.5)	
Don't know/refused	9.0 (2.4-15.5)	16.5 (0.5-32.4)	5.7 (−0.4 to 11.9)	
Education				
Middle school/ some high school	27.2 (17.8-36.6)	49.5 (30.8-68.1)	17.5 (8.0-27.1)	.01
High school graduate/GED[d]	36.2 (26.1-46.3)	24.6 (9.0-40.1)	41.2 (28.7-53.7)	
Associates/ some college	36.6 (27.4-45.9)	26.0 (11.1-40.8)	41.2 (29.6-52.8)	
Employment status				
Unemployed	14.7 (8.0-21.4)	20.3 (5.3-35.3)	12.3 (5.3-19.2)	.22
Employed	42.5 (32.3-52.6)	37.5 (20.0-55.1)	44.6 (32.2-56.9)	
Retired	18.6 (10.6-26.6)	8.9 (0.5-17.4)	22.8 (12.1-33.5)	
Unable to work	12.5 (5.9-19.1)	20.2 (3.2-37.2)	9.1 (3.8-14.5)	
Other	11.8 (4.9-18.7)	13.0 (0.7-25.3)	11.2 (2.8-19.7)	
Health insurance				
Uninsured	17.1 (9.7-24.5)	14.0 (0.8-27.3)	18.4 (9.3-27.5)	.56
Medicaid only	13.7 (6.9-20.6)	21.5 (5.9-37.1)	10.4 (3.6-17.2)	
Medicare only	10.8 (5.7-16.0)	9.9 (1.4-18.4)	11.2 (4.8-17.7)	
Private/other	40.5 (30.5-50.5)	33.5 (16.9-50.0)	43.5 (31.6-55.4)	
Multiple	17.9 (9.8-25.9)	21.1 (3.6-38.6)	16.5 (7.7-25.2)	
Physician visit in past year (% yes)	82.5 (73.6-91.4)	79.3 (63.7-95.0)	83.9 (72.9-94.8)	.63

Characteristic	Total measured obese[a] (BMI[b]≥30 kg/m^2) % (95% CI)	Undiagnosed obese % (95% CI)	Diagnosed obese % (95% CI)	P [c]
	n=140 50.2 (43.0-57.5)	n=43 30.1 (20.7-39.5)	n=97 69.9 (60.5-79.3)	
Source of regular care (% yes)	96.7 (93.1-1.0)	89.1 (77.3-100)	100 (100-100)	.01

[a] A total of 12 people were missing a body mass index value. These individuals were excluded from the chi-square analysis.

[b] BMI: body mass index.

[c] P value for comparison between undiagnosed and diagnosed obese populations; significant at level P<0.05.

[d] GED: general educational development.

ICT Use, Measured Obesity, and Self-Reported Diagnosis of Comorbid Chronic Conditions

Among people with measured obesity, a self-reported diagnosis of hypertension (46.8%) or diabetes (16.8%) was prevalent. Only 36.8% of people with measured obesity had no diagnosis of diabetes or hypertension; 20.5% of the population had all 3 conditions. Isolated measured obesity was associated with higher rates of health-specific ICT use than measured obesity plus comorbid diabetes and/or hypertension diagnosis (77.1% obesity only vs 47.4% obesity and hypertension or diabetes vs 60.7% obesity, hypertension, and diabetes, P=.04; Table 4). Examining rates of specific health-related activities, a statistically significant difference was only noted for accessing Web-based health benefits information (27.6% obesity only vs 18.0% obesity and hypertension or diabetes vs 4.0% obesity, hypertension, and diabetes, P=.04).

Table 4. Comparing information and communication technology activities by presence of comorbid conditions (results of weighted analysis).

ICT[a] activities (% yes)	Measured obesity only[c] % (95% CI)	Measured obesity and hypertension or diabetes % (95% CI)	Measured obesity, hypertension, and diabetes % (95% CI)	P [b]
	n=44 36.8 (26.6-47.0)	n=68 42.7 (32.8-52.6)	n=28 20.5 (12.4-28.6)	
General ICT use	93.7 (87.2-100)	75.0 (60.3-89.6)	73.4 (53.1-93.7)	.05
Any health-specific use	77.1 (61.4-92.7)	47.4 (33.1-61.8)	60.7 (39.3-82.0)	.04
Seek health info online	67.5 (49.8-85.2)	45.9 (31.7-60.0)	48.4 (26.6-70.3)	.15
Use Web-based resources to avoid asking doctor	30.8 (15.1-46.5)	21.9 (10.5-33.4)	15.6 (−1.2 to 32.4)	.41
Web-based access of health benefit info	27.6 (12.4-42.8)	18.0 (7.1-28.9)	4.0 (−1.8 to 9.8)	.04
Participate in online health support group	15.3 (1.7-28.9)	10.5 (1.1-19.8)	2.0 (−2.0 to 6.0)	.22
Web-based access of health records	16.4 (4.2-28.6)	9.1 (2.0-16.1)	13.7 (−1.3 to 28.7)	.57
Web-based medication purchasing	10.8 (0.78-20.7)	8.1 (1.4-14.7)	12.8 (−2.0 to 27.5)	.81
Web-based communication with providers	13.1 (1.8-24.5)	10.2 (3.0-17.4)	15.7 (−1.0 to 32.4)	.80
Use health-related mobile app	9.7 (0.8-18.6)	11.7 (2.2-21.3)	2.0 (−2.0 to 6.0)	.29

[a] ICT: information and communication technology.

[b] P value for comparison between those with "measured obesity only," "measured obesity and hypertension or diabetes," and "measured obesity, hypertension, and diabetes"; significant at level P<0.05.

[c] A total of 12 people were missing a body mass index value. These individuals were excluded from the chi-square analysis.

XSL•FO
RenderX

Discussion

This study describes ICT use in high-poverty African American and Hispanic communities on Chicago's South Side with a disproportionate burden of obesity and obesity-related diseases and examines the association between obesity and ICT use. To our knowledge, this is the only study to ascertain measured BMI, self-reported obesity diagnoses, and ICT use from the same sample. This design, albeit limited by a relatively small sample size, enabled us to generate three new findings. First, we found that ICT use patterns differed by measured obesity status; people with obesity had statistically significant higher rates of both general ICT and health-specific ICT use compared with people without obesity. Second, among people with measured obesity, a physician's diagnosis of obesity was not associated with higher rates of health-specific ICT use or use of Web-based health-related information sources, but it was associated with a higher rate of using Web-based resources to avoid asking questions of a doctor. Finally, an unexpected association between comorbidity burden and health-specific ICT use was found. The highest rates of ICT use were among people with measured obesity only, as compared with those with measured obesity who reported one or two common comorbidities.

In contrast to the Pew study findings that showed lower ICT use among people with other common chronic conditions [14], our study found obesity to be associated with a higher likelihood of general ICT use. Obesity, for some, is associated with difficulties with mobility and other physical activities—factors both potentially related to increased rates of ICT use [27-30]. The physical limitations imposed by obesity may result in a vicious cycle of restricted mobility, greater use of ICT, increasing weight, and increasing physical limitation. These problems are likely exacerbated in high-poverty communities with limited access to health care and other health-promoting resources (eg, fresh food, safe spaces for exercise) [3]. The new finding of relatively high rates of health-specific ICT use among people with obesity in a high-poverty urban community may just reflect greater use of all types of ICT.

Unlike other chronic conditions (eg, hypertension and diabetes), obesity is an outward-facing condition. The social stigma of obesity may lead individuals to ICT-based resources rather than medical care for their health needs [31]. Our findings of higher use of Web-based information resources to avoid asking doctors, among those with diagnosed obesity, along with past work in other stigmatized health conditions, support this possibility. A national survey of adult Internet users found that users with a stigmatized condition (eg, depression, sexually transmitted diseases) were more likely than those with a less stigmatized condition (eg, diabetes, back pain) to report using the Internet as a health information source and as a tool to communicate with clinicians [32]. Obesity was not included among the stigmatized conditions in this study.

Regardless of the drivers behind increased health-specific ICT use, our study suggests that obesity may be a useful target for health-specific ICT-based interventions. The high prevalence of obesity among residents on Chicago's South Side and in

other high-poverty, minority communities, along with the high rate of health-specific use, indicates an already online population with high health needs and risks. ICT-based resources could potentially not only aid in the self-care and management of obesity but also serve as an entry point to provide information and support for routine preventive care and other important health topics in this population.

Among studies of health-related ICT use, this survey was unique in its combination of ICT use measures with anthropometric measures and assessment of self-reported chronic diseases [14,33]. This design enabled closer study of the association between physician-diagnosed obesity and health-specific ICT use. Counter to our hypothesis, obese people who reported a physician's diagnosis of obesity were not more likely to report health-specific ICT use. This result differs from past evidence demonstrating the effect of physician input on health behaviors. In a randomized controlled trial of adults in a primary care setting, individuals randomized to receiving physicians' advice on quitting smoking, reducing fat consumption, and increasing exercise were more likely to believe these topics were relevant to them and more likely to report attempting to quit smoking and making some dietary changes [34]. The absence of differences in health-specific ICT use by provider diagnosis may again reflect differences in obesity as a condition. Unlike diabetes, dyslipidemia, and hypertension diagnosed by blood test or expert measurement, obesity is easily self-diagnosed making the physician's diagnosis less surprising and, potentially, less important.

Obesity status may also be more subject to individual perceptions than other chronic conditions. Past research has demonstrated that people have difficulty in assessing ideal weight [35]. Beyond that, differences in obesity perception by race and ethnicity are well documented in the literature. Using pooled cross-sectional data from National Health and Nutrition Examination Survey, Dorsey et al [36] observed differences in weight perception when comparing non-Hispanic black adults to non-Hispanic white adults; non-Hispanic black adults with obesity who did not perceive themselves as obese had lower odds of desiring weight loss. Given this, it may be that a physician's obesity diagnosis has a different effect when it is discordant with existing cultural norms and an individual's perception of his or her weight. If that is the case, the need to seek out health resources for obesity, including ICT-based health resources, may be less affected by a provider's diagnosis.

Comorbid diagnoses of diabetes and/or hypertension were not found to be associated with a higher likelihood of health-specific ICT use among individuals with measured obesity. This finding contrasts with findings from a 2007 phone survey of US adults that demonstrated a positive correlation between the number of chronic conditions (did not include obesity) and engagement in selected health-specific ICT activities [37]. However, this study's population was younger (>60% aged less than 50 years), had higher educational attainment, and a much lower prevalence of chronic conditions. Also, the racial and ethnic characteristics of the sample were not described. Individuals with obesity in addition to hypertension and/or diabetes are likely to have more frequent in-person contact with health care providers than individuals with obesity alone and may better understand obesity

as, or relating to exacerbation of, a chronic medical condition. It is possible that among individuals with obesity and fewer comorbid conditions, less regular provider contact may result in more unanswered health questions and unmet needs, motivating more health-specific ICT use.
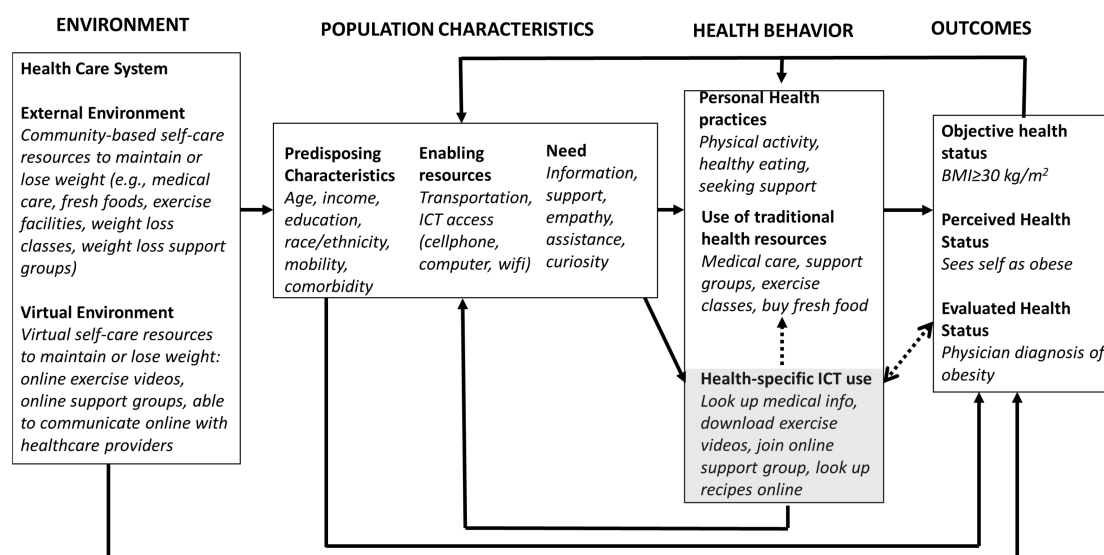
On the basis of the study findings, extant literature, and clinical experience, we propose a preliminary conceptual framework for the relationship between obesity and use of health-specific ICT (Figure 2) [14,20,28,29,31,36,38,39]. The proposed framework is adapted from Andersen's Model of Health Services Utilization [40]; it incorporates obesity as a specific use case and highlights the still incompletely understood interplay between obesity and ICT use. The proposed model also demonstrates the potential for ICT-based services to act as both a complement to traditional health services by enabling access (eg, an individual uses the Internet to find a weight loss support group) and a partial or complete substitute for traditional health services (eg, individual may use the Internet to connect with a weight loss support group online).

The study has several limitations. First, the conservative AAPOR response rate calculation (44.6%) is lower than desired, but it is also consistent with or higher than that reported for other similar surveys in this and other urban populations [41-43]. Of note, Pew reports an 11.6% response rate for its widely cited 2012 phone survey (although used different sampling approach, random digit dialing) [14,44]. Second, although the use of

high-quality population-based probability sampling serves to balance the relatively small sample size and allow for generalization of the findings beyond just the survey respondents, the racial and ethnic characteristics of the studied population may not generalize to other groups. Next, although we were able to report on broad categories of health-specific technology use in this population, we did not collect comprehensive information on exactly how and why people were using health-related ICT. It is also possible that we did not capture all of the current health-specific ICT activities in which people may engage. While we describe an association between measured obesity and health-specific ICT use, we cannot infer causality using cross-sectional data. As has been postulated, it is possible that higher levels of any type of ICT use (including health-specific use) cause sedentariness and increase the likelihood of obesity [29,30].

In this high-poverty urban population, the majority of people with measured obesity reported use of technology for health-specific reasons. This high-risk, already online population presents an opportunity for ICT-based health resources to impact health, especially in communities where the burden of obesity is high. However, understanding current use patterns and potential opportunities for health-specific ICT-based resources is only a first step. The critical next step is evaluating the ability of these technology-based resources to meaningfully impact health care and health outcomes in this high-need, high-risk population.

Figure 2. Proposed conceptual framework for the relationship between obesity and health-specific information and communication technology (ICT) use derived from literature, study results, and clinical experience. Adapted from Andersen's Behavioral Model of Health Services Utilization, the proposed model incorporates obesity as a specific use case. The dashed lines highlight two incompletely understood domains: (1) the relationship between obesity and health-specific ICT use and (2) the potential dual role of health-specific ICT as both an access point to and a replacement for traditional health resources. BMI: body mass index.

Acknowledgments

Authors' Contributions

AG contributed to study conception, design, data analysis and interpretation, and manuscript construction. JM contributed to the study design and execution, data analysis and interpretation, and manuscript content. LBG and VE contributed to the data analysis and interpretation, and manuscript content. RM contributed to study design, data interpretation, and manuscript content. RH and MBW contributed to design of the population health study data collection instrument and protocol, community engagement activities in the preparation and dissemination phases of the study, interpretation of the analytic findings, and review and editing of this manuscript. STL led the research team that administered MAPSCorps and the South Side Health and Vitality Studies population health study data collection. STL also provided material support, contributed to design and execution of the study, and oversaw all aspects of data collection, analysis, and interpretation. She has contributed original writing to and edited drafts of the manuscript for important intellectual content. She has provided material support to the execution of this study.

Conflicts of Interest

RM has grant support from NIH, National Heart, Lung, and Blood Institute (NHLBI) grant K23 (109083) and NHLBI R01 (122457), but this funding was not used to conduct the study. Her time on this project was part of her leadership role with the Robert Wood Johnson Clinical Scholars program. STL is founder and co-owner of a social impact company NowPow, LLC, developed as the sustainable business model expected by a Centers for Medicare & Medicaid Services (CMS) Health Care Innovation Award (1C1CMS330997-03-00, 2012-15). The CMS award did not directly support the research described in this manuscript. The other authors report no conflicts of interest.

Multimedia Appendix 1

Frequency of health-specific information and communication technology activities.

[PDF File (Adobe PDF File), 28KB - jmir_v18i6e182_app1.pdf]

References

1. Centers for Disease Control and Prevention. Adult Obesity Facts URL: http://www.cdc.gov/obesity/data/adult.html [accessed 2016-03-07] [WebCite Cache ID 6fqAbRjNX]
2. National Institute of Diabetes and Digestive and Kidney Diseases. Overweight and Obesity Statistics URL: http://www.niddk.nih.gov/health-information/health-statistics/Pages/overweight-obesity-statistics.aspx [accessed 2016-03-07] [WebCite Cache ID 6fqBQ3VJi]
3. Black JL, Macinko J. The changing distribution and determinants of obesity in the neighborhoods of New York City, 2003-2007. Am J Epidemiol 2010 Apr 1;171(7):765-775 [FREE Full text] [doi: 10.1093/aje/kwp458] [Medline: 20172920]
4. Levine JA. Poverty and obesity in the U.S. Diabetes 2011;60(11):2667-2668. [doi: 10.2337/db11-1118]
5. Chang VW. Racial residential segregation and weight status among US adults. Soc Sci Med 2006 Sep;63(5):1289-1303. [doi: 10.1016/j.socscimed.2006.03.049] [Medline: 16707199]
6. Quinn CC, Clough SS, Minor JM, Lender D, Okafor MC, Gruber-Baldini A. WellDoc mobile diabetes management randomized controlled trial: change in clinical and behavioral outcomes and patient and physician satisfaction. Diabetes Technol Ther 2008 Jun;10(3):160-168. [doi: 10.1089/dia.2008.0283] [Medline: 18473689]
7. Duffey C. 2014 Jul 25. Smartwatches will revolutionise treatment for chronic conditions URL: http://www.theguardian.com/healthcare-network/2014/jul/25/smartwatches-revolutionise-treatment-chronic-conditions [accessed 2016-05-09] [WebCite Cache ID 6hNwYB929]

8.   Dolan B. mobihealthnews. 2013 Jun 13. WellDoc's BlueStar secures first mobile health reimbursement URL: http:/
     /mobihealthnews.com/23026/welldocs-bluestar-secures-first-mobile-health-reimbursement [accessed 2016-05-09] [WebCite
     Cache ID 6hNwhz7Af]

9.   Blumenthal D, Tavenner M. The "meaningful use" regulation for electronic health records. N Engl J Med 2010 Aug
     5;363(6):501-504. [doi: 10.1056/NEJMp1006114]

10. Sequist TD, Taveras EM. Clinic-community linkages for high-value care. N Engl J Med 2014 Dec 4;371(23):2148-2150.
     [doi: 10.1056/NEJMp1408457] [Medline: 25470692]

11. 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]

12. Park M, Kim H, Kim K. Cellular phone and Internet-based individual intervention on blood pressure and obesity in obese
     patients with hypertension. Int J Med Inform 2009 Oct;78(10):704-710. [doi: 10.1016/j.ijmedinf.2009.06.004] [Medline:
     19643661]

13. Schulz PJ, Hulsman RL. Assessing the quality of eHealth programs and their impact on health behaviour and health outcomes
     of users. Patient Educ Couns 2009 Oct;77(1):1-3. [doi: 10.1016/j.pec.2009.08.006]

14. Pew Research Center. 2013 Nov. The Diagnosis Difference URL: http://www.pewinternet.org/files/old-media/
     /Files/Reports/2013/PewResearch_DiagnosisDifference.pdf [accessed 2016-05-08] [WebCite Cache ID 6hNxYPaI7]

15. American Medical Association Annual Meeting. 2013. AMA Adopts New Policies on Second Day of Voting at Annual
     Meeting URL: http://www.ama-assn.org/ama/pub/news/news/2013/2013-06-18-new-ama-policies-annual-meeting.page
     [accessed 2016-05-09] [WebCite Cache ID 6fqGQavoP]

16. The University of Chicago. South Side Health & Vitality Studies URL: https://thestudies.uchicago.edu/node [accessed
     2016-03-07] [WebCite Cache ID 6fqBchMLe]

17. U.S. Census Bureau. 2012. American Community Survey URL: http://factfinder.census.gov [accessed 2016-05-08] [WebCite
     Cache ID 6hNxKENdo]

18. City of Chicago Data Portal. Public Health Statistics--Selected underlying causes of death in Chicago, 2006-2010; URL:
     https://data.cityofchicago.org/Health-Human-Services/Public-Health-Statistics-Selected-underlying-cause/j6cj-r444 [accessed
     2016-05-09] [WebCite Cache ID 6hNxR7L9A]

19. Lindau ST, Makelarski JA, Chin MH, Desautels S, Johnson D, Johnson WE, et al. Building community-engaged health
     research and discovery infrastructure on the South Side of Chicago: science in service to community priorities. Prev Med
     2011;52(3-4):200-207 [FREE Full text] [doi: 10.1016/j.ypmed.2011.01.001] [Medline: 21236295]

20. Rosenstock IM, Strecher VJ, Becker MH. Social learning theory and the Health Belief Model. Health Educ Q
     1988;15(2):175-183. [Medline: 3378902]

21. Lindau ST, James R, Makelarski JA, Sanders E, Johnson D. Comments From the South Side of Chicago on New Haven's
     Inspiring Initiative. Am J Public Health 2012;102(7):e3-e4. [Medline: 22594737]

22. U.S. Census Bureau. United States; 2010. Census Interactive Population Map URL: http://www.census.gov/2010census/
     popmap/ [accessed 2016-03-08] [WebCite Cache ID 6fqGxRNth]

23. Marketing Systems Group. URL: http://www.m-s-g.com/Web/genesys/index.aspx [accessed 2016-03-08] [WebCite Cache
     ID 6fqHEk7FA]

24. Nelson DE, Kreps GL, Hesse BW, Croyle RT, Willis G, Arora NK, et al. The Health Information National Trends Survey
     (HINTS): development, design, and dissemination. J Health Commun 2004;9(5):443-60; discussion 81. [doi:
     10.1080/10810730490504233] [Medline: 15513791]

25. Harris KM, Halpern CT, Whitsel E, Hussey J, Tabor J, Entzel P. The National Longitudinal Study of Adolescent to Adult
     Health. 2009. Research Design General documentation URL: http://www.cpc.unc.edu/projects/addhealth/design [accessed
     2016-05-08] [WebCite Cache ID 6fqHOZjUh]

26. The Association for Public Opinion Research. 2011. Standard Definitions Final Definitions of Case Codes and Outcome
     Rates for Surveys URL: http://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.
     pdf [accessed 2016-05-09] [WebCite Cache ID 6hNyBXFud]

27. Backholer K, Wong E, Freak-Poli R, Walls HL, Peeters A. Increasing body weight and risk of limitations in activities of
     daily living: a systematic review and meta-analysis. Obes Rev 2012 May;13(5):456-468. [doi:
     10.1111/j.1467-789X.2011.00970.x] [Medline: 22212629]

28. Alley DE, Chang VW. The changing relationship of obesity and disability, 1988-2004. JAMA 2007 Nov 7;298(17):2020-2027.
     [doi: 10.1001/jama.298.17.2020] [Medline: 17986695]

29. Chatterjee A, Ross DC. Milken Institute. 2012 Aug. Waistlines of the World: The Effect of Information and Communications
     Technology on Obesity URL: http://www.milkeninstitute.org/publications/view/531 [accessed 2016-05-09] [WebCite
     Cache ID 6hNyNHQOh]

30. 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]

XSL•FO
RenderX

31. Hilbert A, Rief W, Braehler E. Stigmatizing attitudes toward obesity in a representative population-based sample. Obesity (Silver Spring) 2008 Jul;16(7):1529-1534 [FREE Full text] [doi: 10.1038/oby.2008.263] [Medline: 18464749]

32. Berger M, Wagner TH, Baker LC. Internet use and stigmatized illness. Soc Sci Med 2005 Oct;61(8):1821-1827. [doi: 10.1016/j.socscimed.2005.03.025] [Medline: 16029778]

33. Wang JY, Bennett K, Probst J. Subdividing the digital divide: differences in internet access and use among rural residents with medical limitations. J Med Internet Res 2011;13(1):e25 [FREE Full text] [doi: 10.2196/jmir.1534] [Medline: 21371989]

34. Kreuter MW, Chheda SG, Bull FC. How does physician advice influence patient behavior? Evidence for a priming effect. Arch Fam Med 2000 May;9(5):426-433. [Medline: 10810947]

35. Naghshizadian R, Rahnemai-Azar AA, Kella K, Weber MM, Calin ML, Bibi S, et al. Patient perception of ideal body weight and the effect of body mass index. J Obes 2014;2014:491280 [FREE Full text] [doi: 10.1155/2014/491280] [Medline: 25614830]

36. Dorsey RR, Eberhardt MS, Ogden CL. Racial and ethnic differences in weight management behavior by weight perception status. Ethn Dis 2010;20(3):244-250. [Medline: 20828097]

37. Ayers SL, Kronenfeld JJ. Chronic illness and health-seeking information on the Internet. Health (London) 2007 Jul;11(3):327-347. [doi: 10.1177/1363459307077547] [Medline: 17606698]

38. Pew Research Center. Cell Phone and Smartphone Ownership Demographics URL: http://www.pewinternet.org/data-trend/mobile/cell-phone-and-smartphone-ownership-demographics/ [accessed 2016-03-08] [WebCite Cache ID 6fqI7whDT]

39. Pew Research Center. Internet User Demographics URL: http://www.pewinternet.org/data-trend/internet-use/latest-stats/ [accessed 2016-05-08] [WebCite Cache ID 6fqIrVk2y]

40. Aday LA, Andersen R. A framework for the study of access to medical care. Health Services Research 1974;9(3):208-220.

41. Schneider J, Makelarski JA, Van HM, Lipton RB, Abramsohn E, Lauderdale DS, et al. Differential access to digital communication technology: association with health and health survey recruitment within an African-American underserviced urban population. J Urban Health 2011 Jun;88(3):479-492 [FREE Full text] [doi: 10.1007/s11524-010-9533-6] [Medline: 21279451]

42. Margellos-Anast H, Shah AM, Whitman S. Prevalence of obesity among children in six Chicago communities: findings from a health survey. Public Health Reports 2008;123(2):117-125.

43. Thorpe LE, Greene C, Freeman A, Snell E, Rodriguez-Lopez JS, Frankel M, et al. Rationale, design and respondent characteristics of the 2013-2014 New York City Health and Nutrition Examination Survey (NYC HANES 2013-2014). Prev Med Rep 2015;2:580-585 [FREE Full text] [doi: 10.1016/j.pmedr.2015.06.019] [Medline: 26844121]

44. Pew Research Center. 2012 Nov 27. Health Tracking Survey Survey Questions: Revised Topline General documentation URL: http://www.pewinternet.org/files/old-media//Files/Questionnaire/2013/Surveyquestionsandmethodology.pdf [accessed 2016-05-06] [WebCite Cache ID 6hNz8ygtq]

Abbreviations

AAPOR: American Association for Public Opinion Research
BMI: body mass index
ICT: information and communication technology
IRB: Institutional Review Board
SSHVS: South Side Health and Vitality Studies
GED: general educational development

Short Paper

Photoaging Mobile Apps in School-Based Tobacco Prevention: The Mirroring Approach

Titus Josef Brinker[1]; Werner Seeger[1], MD, PhD, Prof. Dr. med.; Fabian Buslaff[1]

Universities of Giessen and Marburg Lung Center, Member of the German Center for Lung Research, Justus-Liebig-University of Giessen, Giessen, Germany

Corresponding Author:
Titus Josef Brinker
Universities of Giessen and Marburg Lung Center
Member of the German Center for Lung Research
Justus-Liebig-University of Giessen
Klinikstr. 33
Giessen, 35392
Germany
Phone: 49 15175084347
Fax: 49 52239474962
Email: titus.brinker@gmail.com

Abstract

Background: Most smokers start smoking during their early adolescence, often with the idea that smoking is glamorous. Adolescent smoking can best be prevented through health education at schools. Interventions that take advantage of the broad availability of mobile phones as well as adolescents' interest in their appearance may be a novel way to improve prevention.

Objective: In this first pilot study, we aimed to use mobile phone technology in accordance with the theory of planned behavior to improve school-based tobacco prevention.

Methods: We used a free photoaging mobile phone app ("Smokerface") in three German secondary schools via a novel method called mirroring. The students' altered three-dimensional selfies on mobile phones or tablets were "mirrored" via a projector in front of their whole grade. Using an anonymous questionnaire, we then measured on a 5-point Likert scale the perceptions of the intervention among 125 students of both genders (average age 12.75 years).

Results: A majority of the students perceived the intervention as fun (77/125, 61.6%), claimed that the intervention motivated them not to smoke (79/125, 63.2%), and stated that they learned new benefits of non-smoking (81/125, 64.8%). Only a minority of students disagreed or fully disagreed that they learned new benefits of non-smoking (16/125, 12.8%) or that they were themselves motivated not to smoke (18/125, 14.4%).

Conclusions: We have presented a novel method to integrate photoaging in school-based tobacco prevention that affects student peer groups and considers the predictors of smoking in accordance with the theory of planned behavior.

Introduction

Most smokers start smoking during their early adolescence, often with a vision in their head of the kind of smoker they want to be [1]. In a recent paper, we introduced photoaging mobile apps that alter a person's self-portrait (ie, a selfie) to predict future appearance [2]. These are considered a novel opportunity for smoking prevention after their effectiveness was first demonstrated by Burford et al [3]. In addition to this, many dermatology publications have called for a novel public health approach in light of new findings on the photoaging effects of smoking [4-8]. A photoaging approach is relevant for teenagers as evidenced by numerous publications demonstrating the influence of attractiveness on self-confidence and quality of life of adolescents [9-11].

According to a recent Cochrane analysis, adolescent smoking can most effectively be prevented by health educators in schools,

XSL•FO
RenderX

but no data for photoaging interventions or their implementation in the school setting are available to date [4].

Methods

To integrate these novel interventions into the school-based setting, we developed and tested the mirroring approach in a pilot study. Mirroring means that the students' altered three-dimensional selfies on mobile phones or tablets were "mirrored" via a projector in front of the whole grade. We included a total sample of 125 Grade 7 students in our cross-sectional pilot study with an average age of 12.75 years (49/125, 39.2% female; 76/125, 60.8% male) attending three secondary schools in Germany. The mirroring approach was implemented by medical students from the Education Against Tobacco non-profit organization [12]. To increase familiarity with the photoaging app (called "Smokerface") and students' participation in the mirroring intervention, students were asked to download the app before our visit, via a letter 3 days in advance. By this means, 52% (21/41) of the grammar school students and 44% (16/36) of the general school students had the photoaging app on their mobile phone when we visited the schools.

In the first 10-minute phase, the displayed face of one student volunteer was used to show the app's altering features to the peer group, providing an incentive for the rest of the class to test the app. Students could interact with their own animated face via touch (sneezing, coughing, etc; see Multimedia Appendix 1). In front of their peers and teachers, they could display their image as a non-smoker/smoker 1, 3, 6, 9, 12, or 15 years in the future (see Figures 1-4). Multiple device displays can be projected simultaneously, which we used to consolidate the altering measures with graphics (eg, to explain wrinkle formation). We implemented mirroring with Galaxy Tab A (Samsung) via Apple's proprietary AirPlay interface using the Android app "Mirroring360" (Splashtop Inc.).

In the second 10-minute phase, students were encouraged to try the app on their own device or one of the tablet computers provided for students who do not own a smartphone or did not download the app. The number of provided tablet computers was calculated so the phase would take up to 10 minutes at the most, factoring in a utilization time of about 2 minutes per student and a rate of 50% of students having downloaded the app. By this calculation, 20 minutes of the mirroring intervention and 10 provided tablets were sufficient to have every student within a grade of 100 pupils successfully photoaged at least once.

The perception of the intervention by the students was measured directly after the intervention via three items in an anonymous survey (see Figure 5) asking (1) "The animation of my 3D-selfie motivates me not to smoke," (2) "I learned new benefits of non-smoking," and (3) "The intervention was fun."

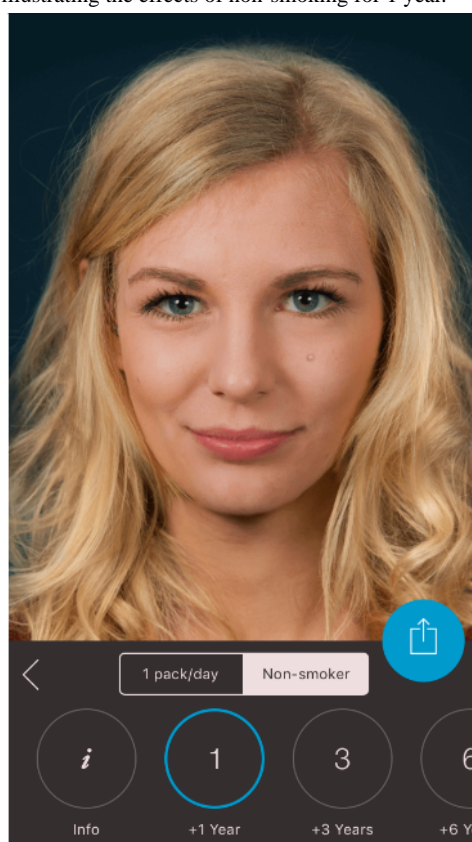**Figure 1.** Screenshot from the Smokerface app illustrating the effects of non-smoking for 1 year.

Figure 2. Screenshot from the Smokerface app illustrating the consequences of smoking a pack a day for 1 year.
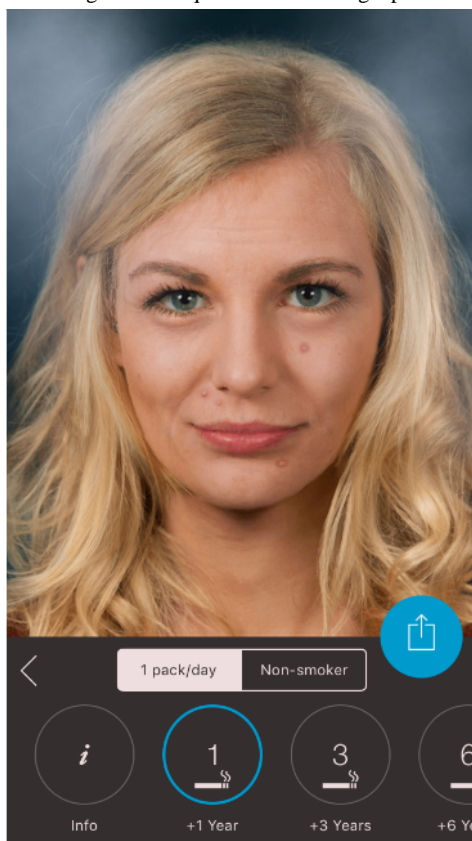

Figure 3. Screenshot from the Smokerface app illustrating the effects of non-smoking for 15 years.
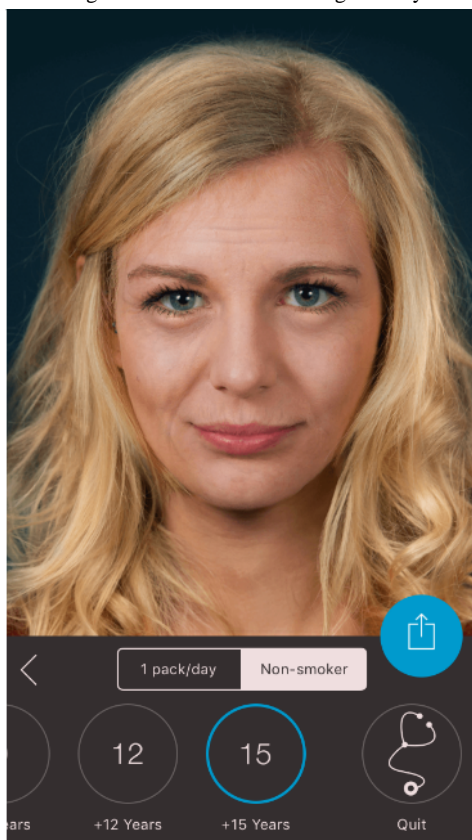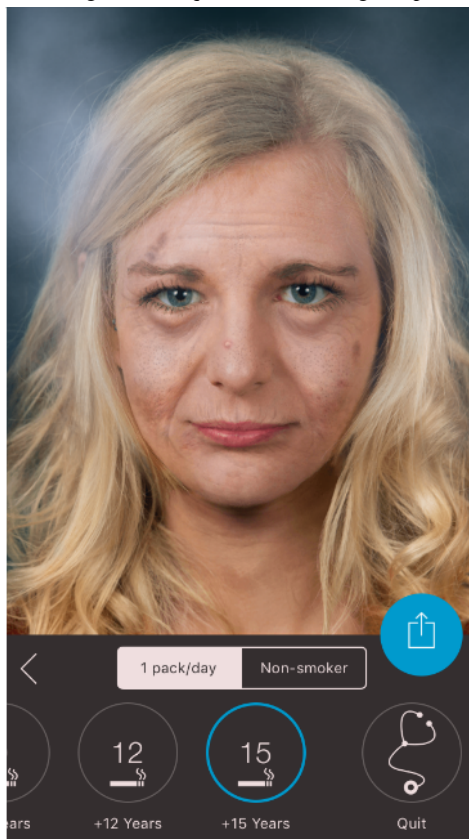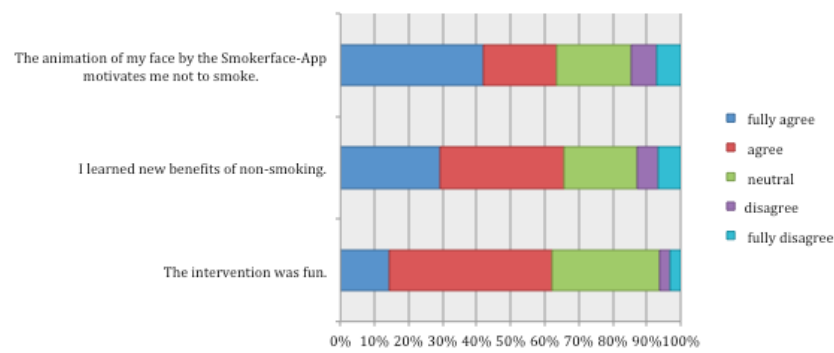
Figure 4. Screenshot from the Smokerface app illustrating the consequences of smoking one pack a day for 15 years.

Figure 5. Main results of the survey.

Results

In our sample, we measured more than 60% (Item 1: 79/125; Item 2: 81/125; Item 3: 77/125) agreement on all three items (see Figure 5). Only a small fraction disagreed or fully disagreed that they learned new benefits of non-smoking (16/125, 12.8%) or that they were themselves motivated not to smoke (18/125, 14.4%). Nearly three-quarters (90/125, 72.0%) of students had already tried the app on their own or a friend's device before the intervention, due to our letter.

The theoretical background of the participant-centered mirroring intervention includes increasing perceived self-efficacy of using the app, which has been proven to encourage repetitive use and is associated with the effectiveness of an intervention according to the theory of planned behavior [13]. Accordingly, 36% (45/124) fully agreed or agreed directly after the intervention that they wanted to use the app again on their own despite the

one-time-use nature of the app and the fact that most of them had used the app at least twice already (neutral=32/124; disagree/fully disagree=47/124; missing=1). By causing direct peer group and teacher reactions to the intervention itself, the subjective norm is affected, which also predicts adolescent smoking [13].

The strongest predictor of adolescent smoking was perceived behavioral control (eg, if students think they could refuse a cigarette successfully) [13]. To this end, an age-appropriate reason not to smoke was integrated into the students' community by means of the app name, "Smokerface". After the intervention, many students would refer to smokers as "smokerfaces" or, when asked why they would not smoke, they stated that they did not want to be a "smokerface." A majority (90%) of the students also agreed that the app was an appropriate tool to convince peers not to smoke when asked after the intervention.

Smoking students motivated to quit were offered free science-based quit advice by the app.

Discussion

Principal Considerations

No evaluation of a tobacco-prevention program using photoaging is available to date despite the fact that mobile phone and tablet use among adolescents is constantly rising both in Germany and globally, providing a novel opportunity for health education [14]. Moreover, authors of dermatology papers as well as those of a recent Cochrane analysis have called for novel public health approaches concerning tobacco prevention [4,5].

The mobile phone app we investigated in this study was easy to implement and was well received. It can be added to existing school-based tobacco-prevention programs.

Limitations

As this study was conducted only in Germany, our results might not be generalizable to other cultural or national settings.

However, cosmetics are used by adolescents in the majority of countries and the World Health Organization is reportedly concerned about the tobacco industry increasingly targeting young females in their advertisements, especially in developing and emerging countries [1]. These developments increase the international relevance of our research.

Conclusion

In summary, we present a novel way of integrating photoaging in school-based tobacco prevention, which affects the students' peer group and considers the predictors of smoking in accordance with the theory of planned behavior. Further research is necessary to evaluate the intervention's effects on the smoking prevalence of adolescents.

A consequence of this pilot study is an integration of the mirroring intervention into our curriculum, which will be evaluated in the Education Against Tobacco Randomized Trial in Germany with approximately 200 secondary schools, 15 medical schools, and around 400 medical students participating.

Authors' Contributions

TJB developed the concept of mirroring for school-based tobacco prevention, designed the study, raised the funds for it to be conducted, wrote the manuscript, conceived of the app, guided its development, and provided guidance with implementation and analysis of the pilot study. WS supported the coordination of the study and proofread the manuscript. FB conducted and analyzed the pilot study, contributed to the design of the intervention, and proofread the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

3D animation of a selfie via the Smokerface app illustrating one of the touch effects (cough) the students can trigger on their own selfies.

[MP4 File (MP4 Video), 556KB - jmir_v18i6e183_app1.mp4]

References

1. World Health Organization. WHO report on the global tobacco epidemic, 2015: Raising taxes on tobacco. 2015. URL: http://apps.who.int/iris/bitstream/10665/178577/1/WHO_NMH_PND_15.5_eng.pdf?ua=1&ua=1 [accessed 2016-06-24] [WebCite Cache ID 6iVhG8E0B]

2. Brinker TJ, Seeger W. Photoaging Mobile Apps: A Novel Opportunity for Smoking Cessation? J Med Internet Res 2015;17(7):e186 [FREE Full text] [doi: 10.2196/jmir.4792] [Medline: 26215210]

3. Burford O, Jiwa M, Carter O, Parsons R, Hendrie D. Internet-based photoaging within Australian pharmacies to promote smoking cessation: randomized controlled trial. J Med Internet Res 2013;15(3):e64 [FREE Full text] [doi: 10.2196/jmir.2337] [Medline: 23531984]

4. Thomas R, McLellan J, Perera R. School-based programmes for preventing smoking. Cochrane Database Syst Rev 2013(4):CD001293. [doi: 10.1002/14651858.CD001293.pub3] [Medline: 23633306]

5. Mosley JG, Gibbs AC. Premature grey hair and hair loss among smokers: a new opportunity for health education? BMJ 1996;313(7072):1616 [FREE Full text] [Medline: 8991008]

6. Schäfer T, Nienhaus A, Vieluf D, Berger J, Ring J. Epidemiology of acne in the general population: the risk of smoking. Br J Dermatol 2001 Jul;145(1):100-104. [Medline: 11453915]

7. Grady D, Ernster V. Does cigarette smoking make you ugly and old? Am J Epidemiol 1992 Apr 15;135(8):839-842. [Medline: 1585896]

8. Okada HC, Alleyne B, Varghai K, Kinder K, Guyuron B. Facial changes caused by smoking: a comparison between smoking and nonsmoking identical twins. Plast Reconstr Surg 2013 Nov;132(5):1085-1092. [doi: 10.1097/PRS.0b013e3182a4c20a] [Medline: 23924651]

9. Dunn LK, O'Neill JL, Feldman SR. Acne in adolescents: quality of life, self-esteem, mood, and psychological disorders. Dermatol Online J 2011;17(1):1 [FREE Full text] [Medline: 21272492]

10. Tasoula E, Gregoriou S, Chalikias J, Lazarou D, Danopoulou I, Katsambas A, et al. The impact of acne vulgaris on quality of life and psychic health in young adolescents in Greece. Results of a population survey. An Bras Dermatol 2012;87(6):862-869 [FREE Full text] [Medline: 23197205]

11. Reljić V, Maksimović N, Janković J, Mijović B, Perić J, Janković S. Evaluation of the quality of life in adolescents with acne. Vojnosanit Pregl 2014 Jul;71(7):634-638. [Medline: 25109109]

12. Brinker T, Stamm-Balderjahn S, Seeger W, Klingelhöfer D, Groneberg DA. Education Against Tobacco (EAT): a quasi-experimental prospective evaluation of a multinational medical-student-delivered smoking prevention programme for secondary schools in Germany. BMJ Open 2015;5(9):e008093 [FREE Full text] [doi: 10.1136/bmjopen-2015-008093] [Medline: 26384722]

13. McEachan R, Conner M, Taylor N, Lawton R. Prospective prediction of health-related behaviours with the Theory of Planned Behaviour: a meta-analysis. Health Psychology Review 2011 Sep;5(2):97-144. [doi: 10.1080/17437199.2010.521684]

14. Kempf D, Holdampf-Wendel A. BITKOM. 2014. Kinder und Jugend 3.0 URL: https://www.bitkom.org/Presse/Anhaenge-an-PIs/2014/April/BITKOM-PK-Kinder-und-Jugend-3-0.pdf [accessed 2016-06-24] [WebCite Cache ID 6iVbRNs2o]

XSL•FO

RenderX

The Sydney West Knowledge Portal: Evaluating the Growth of a Knowledge Portal to Support Translational Research

Anna Janssen[1*], BA, MA; Tracy Elizabeth Robinson[2*], BA(Hons), RN (Registered Nurse); Pamela Provan[3*], BSc; Tim Shaw[1*], BSc(Hons), PhD

[1]Research in Implementation Science and eHealth, Faculty of Health Sciences, The University of Sydney, Sydney University, Australia

[2]Faculty of Health, Discipline of Nursing and Midwifery, University of Canberra, Canberra, Australia

[3]Centre for Cancer Research, The Westmead Institute for Medical Research, Sydney, Australia

[*]all authors contributed equally

Corresponding Author:
Anna Janssen, BA, MA
Research in Implementation Science and eHealth
Faculty of Health Sciences
The University of Sydney
Level 2, Charles Perkins Centre D17
Sydney University, 2006
Australia
Phone: 61 9036 9406
Fax: 61 9351 4561
Email: anna.janssen@sydney.edu.au

Abstract

Background: The Sydney West Translational Cancer Research Centre is an organization funded to build capacity for translational research in cancer. Translational research is essential for ensuring the integration of best available evidence into practice and for improving patient outcomes. However, there is a low level of awareness regarding what it is and how to conduct it optimally. One solution to addressing this gap is the design and deployment of web-based knowledge portals to disseminate new knowledge and engage with and connect dispersed networks of researchers. A knowledge portal is an web-based platform for increasing knowledge dissemination and management in a specialized area.

Objective: To measure the design and growth of an web-based knowledge portal for increasing individual awareness of translational research and to build organizational capacity for the delivery of translational research projects in cancer.

Methods: An adaptive methodology was used to capture the design and growth of an web-based knowledge portal in cancer. This involved stakeholder consultations to inform initial design of the portal. Once the portal was live, site analytics were reviewed to evaluate member usage of the portal and to measure growth in membership.

Results: Knowledge portal membership grew consistently for the first 18 months after deployment, before leveling out. Analysis of site metrics revealed members were most likely to visit portal pages with community-generated content, particularly pages with a focus on translational research. This was closely followed by pages that disseminated educational material about translational research.

Conclusions: Preliminary data from this study suggest that knowledge portals may be beneficial tools for translating new evidence and fostering an environment of communication and collaboration.

(J Med Internet Res 2016;18(6):e170)   doi:10.2196/jmir.5786

Introduction

The term translational research is widely used in the health sector. It describes a spectrum of research from translating basic sciences into new treatment options for patients to improving quality of care by changing patient and health professional behaviors [1].

In spite of the widespread use of the term translational research, there is a low level of awareness regarding what it is and how to conduct it. This knowledge gap compromises the ability of health professionals to design effective translational research studies. The important role of translational research for translating evidence into practice to improve quality of care for patients makes this gap a cause for concern [2].

To address this gap the Sydney West Translational Cancer Research Centre (SW-TCRC) built an web-based knowledge portal: The Sydney West Knowledge Portal (SWKP). The SW-TCRC is a network of cancer care professionals and researchers in Western Sydney, New South Wales, Australia. The portal is a resource to support growth of the SW-TCRC and connect members who are dispersed by both geography and discipline. Additionally, it aims to disseminate key information to members in order to raise awareness and understanding of translational research and reduce the knowledge gap.

The SWKP is underpinned by the Westfall model of translational research. This model identifies 3 phases of research translation: T1, T2, and T3. T1 describes the translation of research from the laboratory to humans, T2 describes the translation to patients, and T3 describes the translation to practice [3]. This model is widely cited in the literature [4,5] and is also endorsed by the Cancer Institute NSW. The development of the SWKP was also informed by literature on the development of health professional networks [6] and strategies for optimizing knowledge translation [7].

A knowledge portal is a platform that drives knowledge production, integration, and management [8]. A distinction is made between knowledge and information. Knowledge represents what an individual knows. Information is knowledge that organizations and individuals can utilize in a meaningful manner. Although knowledge portals originated in the information technology sphere, they have since been more broadly adopted. In the health sector, they have been used for disseminating knowledge about quality improvement [9] and for improving patient health literacy [10]. There is also growing research interest into their use of tools to disseminate knowledge about translational research [11,12]. Knowledge portals are able to disseminate knowledge to a dispersed member base. They also have potential for increasing awareness and understanding of best practice in research translation.

This short report describes the preliminary design and growth of the SWKP, including an evaluation methodology to measure its reach and impact.

Methods

An adaptive methodology was used to capture the design, growth, and evaluation phases of the SWKP. Adaptive methods modify aspects of the research in response to data received during the study, which is important when evaluating in a multiphase intervention like the SWKP.

The Development Phase

Stakeholder consultation was undertaken with key individuals in the SW-TCRC. This process identified goals and boundaries of the portal by identifying optimal membership and member needs. The process also helped to identify the organizational objectives for the portal.

Data from the consultation process were analyzed and used to inform a portal schema. This schema underpinned the design of the portal and was used as a benchmark to evaluate how the portal evolved during its growth phase.

The Growth Phase

This phase began when the portal was made publicly accessible. Membership of the portal was free to all members of the SW-TCRC network. Additionally, exclusive research support opportunities were promoted through the SWKP to encourage individuals to sign up. Word-of-mouth referral from cancer researchers was the primary means of recruitment for the portal. An email promotion campaign was also used to engage leaders within Western Sydney to encourage them to refer members. Finally, a grassroots advertising campaign was used to promote the portal across 3 hospitals within Western Sydney and at conferences and other scientific meetings.

The number of portal access requests made through the web-based sign-up form was analyzed to assess membership growth. Data from the application form were also used to determine distribution of members across the translational pipeline and their research focus. This form also provided referral information for each new member, allowing the research team to identify the portaling the research team to identify were used to determine member behaviors around the site and to identify the resources that attracted most traffic.

Results

The Development Phase

The development phase took 12 months to complete. At the start of this phase, a 3-month stakeholder consultation was undertaken. This included meetings with the management of the SW-TCRC, representatives of the organization's Education and Training Sub-Committee, and a range of potential members including administrators, medical oncologists, nurses, and cancer researchers.

Meeting minutes from consultation sessions were analyzed and used to develop a schema document. This document proposed designing the portal around 3 pillars: Information, Community, and Education. Information would cover content on events, funding, and news specific to the SW-TCRC. Community would incorporate resources to encourage research collaborations or to share personalized content such as individual research projects. Education would cover content developed to transfer knowledge on research translation to support growth of research in this area.

The schema also identified features to incorporate into the portal design during the development phase. This included an activities section, discussion capabilities for member-organization interaction and member-member interaction, member blogs, and site notifications and newsletters to keep members connected.

The portal was built using the Refinery content management system (CMS). The CMS was chosen because it is free to use and adapt, meaning it could be customized to suit the needs of the SWKP. Additionally, it had a minimalistic administration interface, which made managing site content easier for SW-TCRC staff.

The Growth Phase

The SWKP launched in February 2013. In August 2013, it reached its target of 200 members and was continuing to grow. By the end of 2013, the number of portal members had grown to 382. However, in the second half of 2014 membership growth began to slow down, and by early 2015 membership appeared to have equalized at 399 members.

Once the SWKP launched, governance was provided by the SW-TCRC Education and Training Sub-Committee, who reviewed the growth of the portal quarterly. Development of content for the SWKP was undertaken by a small team of editors within the SW-TCRC, under the guidance of implementation science and translational research experts. Additionally, SWKP members were given opportunities to contribute to the portal. One example of this was the publication of member-generated conference reports. Between February 2013 and December 2014, a total of 17 conference reports were published.

Analysis of member referral data showed that 21% (n=86) of current members referred new members. Of these, 50% (42/86) referred 1 new member, 21% (18/86) referred 2 new members, 16% (14/86) referred 3 new members, and 14% (12/86) referred 4 or more. Of the top 14% of referrers, 75% (9/12) were in leadership positions: 44% (4/9) referrers were directors, 33% (3/9) were managers, and 22% (2/9) were department heads.

Website analytics covering the first year and a half after portal deployment showed there were 7808 page views across 1154 sessions. The average session duration was 11 minutes and 6 seconds and consisted of 6 page views across the site. The majority of visits, 95.9% (1107/1154), were by returning visitors.

The top 50 pages across the 3 categories Information, Community, and Education were also reviewed. This showed that Community pages were the most popular, making up 27 of the top 50 visited pages on the portal (54%). This was followed by Education, which represented 9 of the 50 most visited portal pages (18%), and Information, which made up 5 of the 50 most visited pages (10%). A remaining 9 pages were classified as "other" (18%) and included pages such as "About SW-TCRC."

Discussion

Developing and sustaining knowledge portals is a significant challenge for organizations [13]. However, results of this study show that health professionals are interested in participating in such web-based communities. This is consistent with existing literature on the role of online research networks in health care [10,14,15]. Additionally, this study emphasizes the importance of the development phase in ensuring relevance and sustainability of web-based knowledge portals. A year was dedicated to designing the SWKP, which included development of key documents to clarify the technological and social structure that underpinned the community.

This paper also shows the important role of community members in growing and sustaining an web-based portal. In this study, members acted as content developers and as community recruiters. This finding contributes to the current literature on how knowledge portals act as foundations from which other web-based communities such as communities of practice and online collaborative networks grow [15].

Finally, SWKP metrics show that community-focused content is more appealing to members than other content. It was more than twice as popular as Education content and 5 times more popular than content in the Information category. Encouraging collaborations between health professionals is known to be challenging, and there is a lack of research into building and sustaining tools to do so [16]. SWKP data suggest that knowledge portals may be beneficial tools for fostering an environment of communication and collaboration.

Conflicts of Interest

None declared.

References

1. Woolf SH. The meaning of translational research and why it matters. JAMA 2008 Jan 9;299(2):211-213. [doi: 10.1001/jama.2007.26] [Medline: 18182604]
2. Dougherty D, Conway PH. The "3T's" road map to transform US health care: the "how" of high-quality care. JAMA 2008 May 21;299(19):2319-2321. [doi: 10.1001/jama.299.19.2319] [Medline: 18492974]
3. Westfall JM, Mold J, Fagnan L. Practice-based research: "Blue Highways" on the NIH roadmap. JAMA 2007 Jan 24;297(4):403-406. [doi: 10.1001/jama.297.4.403] [Medline: 17244837]

4.   Drolet BC, Lorenzi NM. Translational research: understanding the continuum from bench to bedside. Transl Res 2011 Jan;157(1):1-5. [doi: 10.1016/j.trsl.2010.10.002] [Medline: 21146144]

5.   Rubio DM, Schoenbaum EE, Lee LS, Schteingart DE, Marantz PR, Marantz KE, et al. Defining Translational Research: Implications for Training. Academic medicine: journal of the Association of American Medical Colleges 2010;85(3):470-475. [doi: 10.1097/ACM.0b013e3181ccd618]

6.   Cunningham FC, Ranmuthugala G, Plumb J, Georgiou A, Westbrook JI, Braithwaite J. Health professional networks as a vector for improving healthcare quality and safety: a systematic review. BMJ Qual Saf 2012 Mar;21(3):239-249 [FREE Full text] [doi: 10.1136/bmjqs-2011-000187] [Medline: 22129933]

7.   Armstrong R, Waters E, Dobbins M, Lavis JN, Petticrew M, Christensen R. Knowledge translation strategies for facilitating evidence-informed public health decision making among managers and policy-makers. Cochrane Database Syst Rev 2011(6):CD009181. [doi: 10.1002/14651858.CD009181]

8.   Firestone JM. Enterprise knowledge portals: what they are and what they do. Knowledge and Innovation: Journal of the KMCI 2000;1(1):85-108 [FREE Full text]

9.   Wardak D, Parker P, Shaw T. ISQua Knowledge: fostering a sense of community for the emerging field of quality in health care. In: Carvalho L, Goodyear P, editors. The Architecture of Productive Learning Networks. New York: Routledge; 2014:181-192.

10. Sarkar U, Karter AJ, Liu JY, Adler NE, Nguyen R, Lopez A, et al. The literacy divide: health literacy and the use of an internet-based patient portal in an integrated health system-results from the diabetes study of northern California (DISTANCE). J Health Commun 2010;15 Suppl 2:183-196 [FREE Full text] [doi: 10.1080/10810730.2010.499988] [Medline: 20845203]

11. Archambault PM, Turgeon AF, Witteman HO, Lauzier F, Moore L, Lamontagne F, Canadian Critical Care Trials Group. Implementation and Evaluation of a Wiki Involving Multiple Stakeholders Including Patients in the Promotion of Best Practices in Trauma Care: The WikiTrauma Interrupted Time Series Protocol. JMIR Res Protoc 2015;4(1):e21. [doi: 10.2196/resprot.4024] [Medline: 25699546]

12. Mathew D, McKibbon KA, Lokker C, Colquhoun H. Engaging with a Wiki related to knowledge translation: a survey of WhatisKT Wiki users. J Med Internet Res 2014;16(1):e21 [FREE Full text] [doi: 10.2196/jmir.3001] [Medline: 24449712]

13. Benbya H, Passiante GN, Belbaly NA. Corporate portal: a tool for knowledge management synchronization. International Journal of Information Management 2004 Jun;24(3):201-220 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2003.12.012]

14. Brulet A, Llorca G, Letrilliart L. Medical wikis dedicated to clinical practice: a systematic review. J Med Internet Res 2015;17(2):e48 [FREE Full text] [Medline: 25700482]

15. Robinson T, Rankin N, Janssen A, Mcgregor D, Grieve S, Shaw T. Collaborative research networks in health: a pragmatic scoping study for the development of an imaging network. Health Res Policy Syst 2015;13:76 [FREE Full text] [doi: 10.1186/s12961-015-0067-y] [Medline: 26652643]

16. van Baalen P, Bloemhof-Ruwaard J, van Heck E. Knowledge Sharing in an Emerging Network of Practice: The Role of a Knowledge Portal. European Management Journal 2005 Jun;23(3):300-314. [doi: 10.1016/j.emj.2005.04.008]

Abbreviations

CMS: content management system
SW-TCRC: Sydney West Translational Cancer Research Centre
SWKP: Sydney West Knowledge Portal

XSL•FO
RenderX

[XSL•FO]

RenderX

Original Paper

Manipulating Google's Knowledge Graph Box to Counter Biased Information Processing During an Online Search on Vaccination: Application of a Technological Debiasing Strategy

Ramona Ludolph[1], MPH; Ahmed Allam[2], PhD; Peter J Schulz[1], PhD

[1]Institute of Communication and Health, Faculty of Communication Sciences, University of Lugano (Università della Svizzera italiana), Lugano, Switzerland

[2]Department of Pathology, Yale University School of Medicine, New Haven, CT, United States

Corresponding Author:
Ramona Ludolph, MPH
Institute of Communication and Health
Faculty of Communication Sciences
University of Lugano (Università della Svizzera italiana)
Via G. Buffi 13
Lugano, 6904
Switzerland
Phone: 41 58666 ext 4821
Fax: 41 586664647
Email: ramona.alexandra.ludolph@usi.ch

Abstract

Background: One of people's major motives for going online is the search for health-related information. Most consumers start their search with a general search engine but are unaware of the fact that its sorting and ranking criteria do not mirror information quality. This misconception can lead to distorted search outcomes, especially when the information processing is characterized by heuristic principles and resulting cognitive biases instead of a systematic elaboration. As vaccination opponents are vocal on the Web, the chance of encountering their non evidence-based views on immunization is high. Therefore, biased information processing in this context can cause subsequent impaired judgment and decision making. A technological debiasing strategy could counter this by changing people's search environment.

Objective: This study aims at testing a technological debiasing strategy to reduce the negative effects of biased information processing when using a general search engine on people's vaccination-related knowledge and attitudes. This strategy is to manipulate the content of Google's knowledge graph box, which is integrated in the search interface and provides basic information about the search topic.

Methods: A full 3x2 factorial, posttest-only design was employed with availability of basic factual information (comprehensible vs hardly comprehensible vs not present) as the first factor and a warning message as the second factor of experimental manipulation. Outcome variables were the evaluation of the knowledge graph box, vaccination-related knowledge, as well as beliefs and attitudes toward vaccination, as represented by three latent variables emerged from an exploratory factor analysis.

Results: Two-way analysis of variance revealed a significant main effect of availability of basic information in the knowledge graph box on participants' vaccination knowledge scores ($F_{2,273}=4.86$, $P=.01$), skepticism/fear of vaccination side effects ($F_{2,273}=3.5$, $P=.03$), and perceived information quality ($F_{2,273}=3.73$, $P=.02$). More specifically, respondents receiving comprehensible information appeared to be more knowledgeable, less skeptical of vaccination, and more critical of information quality compared to participants exposed to hardly comprehensible information. Although, there was no significant interaction effect between the availability of information and the presence of the warning, there was a dominant pattern in which the presence of the warning appeared to have a positive influence on the group receiving comprehensible information while the opposite was true for the groups exposed to hardly comprehensible information and no information at all. Participants evaluated the knowledge graph box as moderately to highly useful, with no significant differences among the experimental groups.

Conclusion: Overall, the results suggest that comprehensible information in the knowledge graph box positively affects participants' vaccination-related knowledge and attitudes. A small change in the content retrieval procedure currently used by

Google could already make a valuable difference in the pursuit of an unbiased online information search. Further research is needed to gain insights into the knowledge graph box's entire potential.

Introduction

Background

There is little doubt today that the Internet has revolutionized consumers' information search [1]. One of people's major motives for going online is the search for health-related information [2,3]. However, the Web does not only offer the opportunity to access an abundance of information from a broad variety of sources but also bears the peril of "getting lost in information" [4] and an increased risk of encountering misinformation [1]. This holds especially true for the topic of vaccination. Despite the large evidence base that proves vaccines' safety and effectiveness [5-8], vaccination opponents continue spreading numerous myths online about immunization [9-11]. Their high visibility on the Web [10,12,13] might have contributed to vaccination being a controversially discussed topic with many people being concerned about potential side effects and hesitant to adhere to official vaccination recommendations [14-19].

General search engines play a major role when it comes to the visibility of a topic. Given that most people start their search with a general search engine like Google, Bing, or Yahoo [2,20,21], a search engine's sorting and ranking criteria can directly influence the search outcomes [22-24]. In turn, websites can be designed to better meet those criteria and achieve higher visibility [25]. Yet, most consumers do not seem to be aware of the logic behind search engines' retrieval algorithms [23,24]. An eye tracking experiment revealed, for instance, that participants trusted Google to rank search results according to their pertinence to the search query [24]. This led to a closer scrutiny of highly positioned search results, regardless of their actual relevance [24]. Another study found the sorting and ranking criteria to impact information seekers' attitudes and knowledge [23]. More specifically, when the authors manipulated the ratio of pro- and antivaccination websites displayed by Google, a detrimental effect of a high share of antivaccination websites was detected [23]. In contrast, the complete absence of negative information resulted in higher knowledge and more favorable attitudes toward vaccination [23]. The high share of misleading or false information concerning vaccination on the Web thus bears a peril for health information seekers [11], especially if the websites fulfill a search engine's conditions to get a high rank on the results pages.

One explanation for consumers' neglect of the logic behind the results displayed by a search engine and the associated suboptimal search outcomes could be a lack of systematic information processing. Instead of systematically elaborating on the search and its results, online information seekers seem to apply heuristic principles [24,26-28]. Whereas systematic information processing is characterized by a "thorough, in-depth, complete, and well-advised processing of all given information," heuristic processing can be described as "relying on cues that signal truth, quality, or validity" [28]. Especially experienced Web users, as compared to inexperienced ones, were found to process information rather heuristically when using a search engine [28]. A heuristic can be defined as a "cognitive shortcut that relies on little information and modest cognitive resources" and often results in satisfying outcomes [29]. So-called fast and frugal heuristics exploit the composition of a certain environment in a prompt and relatively effortless manner [30] and might thus help experienced information seekers to quickly scan a search engine's results without investing too much cognitive effort [24].

However, the ignorance of crucial parts of information can also result in cognitive biases [29,31,32], a phenomenon that was primarily demonstrated in the classic experiments by Tversky and Kahneman [31]. Cognitive biases are understood as "systematic error[s] in judgment and decision-making…which can be due to cognitive limitations, motivational factors, and/or adaptations to natural environments" [29]. In the context of online information seeking via a general search engine, the use of heuristic principles and accompanied reduction of complexity might lead to derogated search outcomes such as a biased processing or interpretation of the results [23,24]. Eventually, this can have negative effects on people's subsequent judgment and decision making if it is based on the search results [23].

Previous Studies

Indeed, several studies report the occurrence and detrimental effects of search-related cognitive biases [23,33-35]. Findings of one study demonstrated that clinicians experience cognitive biases such as anchoring, exposure, or order bias during the online information search [33]. More specifically, participants' prior beliefs about the search topic, the mere time they spent on processing the information, and the position where the information was placed in the retrieval system influenced people's postsearch decisions [33]. Moreover, it was found that health information seekers who start their search with a "strong specific hypothesis" about the search topic tend to focus their search on the verification of this belief and are inclined to interpret the search results as supporting their initial idea [34]. These tendencies are also known as positive hypothesis testing and confirmation bias and were supported by another study. There, the author demonstrated that participants "were very prone to positive hypothesis testing when they searched for health information using a popular search engine" [35].

To overcome these unwanted effects of cognitive biases, debiasing interventions were designed and tested. Debiasing techniques aim at eliminating, reducing, or reversing detrimental effects on judgment and decision making caused by cognitive biases [36-40]. Taking into account the underlying causes for the occurrence of those biases, one can differentiate between motivational, cognitive, and technological debiasing techniques [39]. Motivational strategies involve the promise of incentives or holding people accountable for their judgment in order to increase their motivation [39]. Cognitive and technological strategies are based on the assumption that intuitively applied decision strategies are imperfect, "but that they can be replaced by strategies that approach normative standards" [39]. For instance, "consider the opposite" could be regarded as a cognitive strategy [39]. Technological strategies, in turn, comprise the provision of external tools and techniques to improve the decision environment [39].

The application of technological debiasing strategies seems to be especially promising in the context of online health information seeking since the Web setting offers new opportunities to integrate those external tools [37]. This renders a long and resource-intense debiasing of the individual unnecessary but allows for a central change of the information seekers' environment. Although the emphasis on the importance of an individual's environment was mainly stimulated by the fast and frugal heuristics program [29], it also initiated the development of new technological debiasing approaches. One example is the implementation of a search interface that urges the user to order the search results in a particular manner [41]. This change of the search environment was shown to have a debiasing effect [41]. Another author tested two further technological debiasing techniques in the context of online health information seeking, namely recommendation and incorporation [35]. The recommendation strategy implied the automated suggestion of additional information by the search engine based on the entered keywords [35]. However, this strategy did not lead to a significant reduction of participants' positive hypothesis testing [35]. In contrast, the incorporation strategy turned out to be an effective debiasing technique. It involved the configuration of a search engine that automatically built in search results with an opposite meaning than the ones entered by the participants [35]. For instance, when consumers entered "hypertension," search results referring to "low blood pressure" were also displayed [35].

In sum, implemented technological debiasing interventions either involved the design of a specific search engine, which is complex and time consuming [41], or were not successful in reducing the undesired effects [35], or rather "outwitted" the information seekers instead of transparently countering their biased information processing [35]. Hence, the existing debiasing attempts call for new ideas and further research in implementing technological debiasing strategies. Drawing on the assumed potential of changing environmental structures, this study seeks to test the debiasing effects of a knowledge graph box. A knowledge graph box is a little box containing a short summary related to the search topic, displayed on the right upper side of the screen after the keywords of the search are entered. It was introduced by Google in mid-2012 [42] with the goal of providing semantic information related to the search topic gathered from various knowledge bases (eg, Freebase/Wikidata). In our study, the knowledge graph box was visible to participants occupied with an information search task as part of an online experiment intended to mitigate people's deficient processing of health information found via Google.

Objective and Hypotheses

Our aim was to design a debiasing intervention that would interrupt information seekers' biased information processing and lead to a more systematic scrutinizing of search results, finally resulting in more knowledge and better attitudes toward vaccination. The design of the intervention was based on two lines of research. First, we followed the assumptions of technological debiasing by changing the search environment. More specifically, we added a manipulated version of Google's knowledge graph box to the search interface. Second, the knowledge graph box's content was designed in accordance with the assumption that a message's effectiveness should mainly depend on its content, as opposed to receiver or source characteristics [43].

The content was divided into two parts that could possibly interfere with the biased information search, namely making (hardly) comprehensible basic factual information about vaccination available and a warning message. In general, we assumed the content of the knowledge graph box would serve as an implicit reference for the subsequently explored search results.

Concerning the basic factual information, two definitions of vaccination were provided. These were neutral in tone to avoid an unintended biasing of participants or reactance. However, they represented a different level of comprehensibility. Several experiments in the context of medical information revealed that comprehensible scientific content is more persuasive and receives stronger agreement from laypeople than does incomprehensible content [44,45]. Comprehensibility refers to the extent to which a health information seeker would quickly and easily understand what they read (see also [44]) and was operationalized in two ways. The first indicator for comprehensibility was the readability of the text as assessed by a readability formula such as the Flesch-Kincaid Reading Ease, where higher scores point to easier readability [46,47]. For the comprehensible definition, the formula provided a score of 23.7, whereas the hardly comprehensible one received a score of only -12.1 [47]. The second indicator for comprehensibility was the number of unexplained technical terms [44]. The hardly comprehensible definition used more unexplained medical terms such as "antigenic material," "adaptive immunity," or "pathogen," whereas the comprehensible definition linked common language explanations to the medical expressions. Extending the results of previous research on the persuasive effects of comprehensibility to our context, we expected participants exposed to the comprehensible version of the basic information to show a higher postsearch vaccination knowledge level (Hypothesis 1) and more favorable attitudes toward vaccination (Hypothesis 2) as compared to the groups receiving hardly comprehensible information or no basic information at all.

As a second feature of the knowledge graph box, a warning message was intended to serve as a quality alert to the health information seekers. This is based on research from the area of consumer psychology where participants were found to detect manipulative practices more likely after they had been explicitly warned about their occurrence [48]. To avoid unintended reactions on the participants' side, we refrained from an explicit warning about vaccination opponents' views. Instead, a neutral warning about the existence of false or misleading information about vaccination on the Web was used. This quality alert was predicted to interrupt people's heuristic use of the search engine since it is a new and unexpected search element [28,49] pointing to the need of a thorough inquiry of the search results [48]. The subsequent shift to a more systematic search should then lead to an elaborate information processing characterized by a closer scrutinizing of the search results' quality. Accordingly, we hypothesize that the warning in the knowledge graph box will eventually be associated with higher postsearch knowledge levels (Hypothesis 3) and a more positive attitude toward vaccination (Hypothesis 4).

To achieve a comprehensive testing of the debiasing intervention, the two content-related factors were also combined. Here, we assumed that there could be an interaction effect of the availability of basic information (comprehensible vs incomprehensible) with the warning. However, as we could only speculate about its direction, we formulate Research Question 1: How do basic information and warning message interact to affect consumers' vaccination-related knowledge and attitudes?

As this is, to our knowledge, the first study that investigates the impact of an experimentally manipulated knowledge graph box

integrated into the search process, we are also interested in consumers' perception and evaluation of it. Therefore, Research Question 2 asks: How is the knowledge graph box perceived and evaluated by the participants?

Methods

Design

In order to answer the research questions and test the hypotheses, a posttest-only online experiment was conducted. A full 3x2 factorial design was employed with the content of the knowledge graph box being varied. The first factor was availability of basic factual information on vaccination in the form of an evidence-based, medical definition (comprehensible vs hardly comprehensible vs not present) and the second factor was the presence or absence of a warning message about the occurrence of false information.

There were two versions of basic information that were extracted from two different sources, namely the World Health Organization (WHO) for the comprehensible version and Wikipedia for the less comprehensible version.

The warning stated that vaccination was a controversially discussed topic and that one would encounter false or misleading information about it on the Web. The warning did not support or elicit a standpoint toward the searched health topic.

In sum, the manipulations of the knowledge graph box resulted in the comparison of six different groups (see Table 1). The study was approved by the Ethical Committee of the University of Lugano, Switzerland. Further, all participants were asked for informed consent before taking part in the experiment.

Table 1. Experimental design and group allocation.

Knowledge graph box manipulation	Availability of basic information		
	Comprehensible basic information (WHO)	Hardly comprehensible basic information (Wikipedia)	No basic information
Warning of the presence of false information			
No warning	Group 1	Group 2	Group 6 (Control group)
Warning present	Group 4	Group 5	Group 3

Manipulation of the Knowledge Graph Box

As we aimed for realistic results, the manipulation of the knowledge graph box mirrored the current practice of Google. It retrieves the semantic information from knowledge bases such as Wikidata [42], which acts as "central storage of structured data for Wikipedia" [50]. Hence, when searching for "vaccination" (as of January 2016), the knowledge graph box would display the same semantic information in a box right next to the search results that is also found in the corresponding Wikipedia article (see Multimedia Appendix 1 for a demonstration). Moreover, the knowledge graph box included

a link to the original articles where the extracted information can be found.

In addition to the manipulation of the knowledge graph box, all groups were exposed to the traditional search results as displayed in the Google search template (Figure 1). The knowledge graph box (displaying basic factual information and/or the warning) as well as the search results page implemented in this experiment used the same template (eg, design, fonts, colors) as provided by Google. Figure 2 depicts the template used showing the manipulated experimental factors (different versions of basic information and the warning message) as well as their corresponding experimental groups as mentioned in Table 1.

Figure 1. Search result template mimicking Google's search used during the experiment.

Figure 2. Three examples of the knowledge graph box as it appeared on the screen during the search: (A) Group 1 (comprehensible information from WHO), (B) Group 2 (hardly comprehensible information from Wikipedia), (C) Group 3 (warning message).
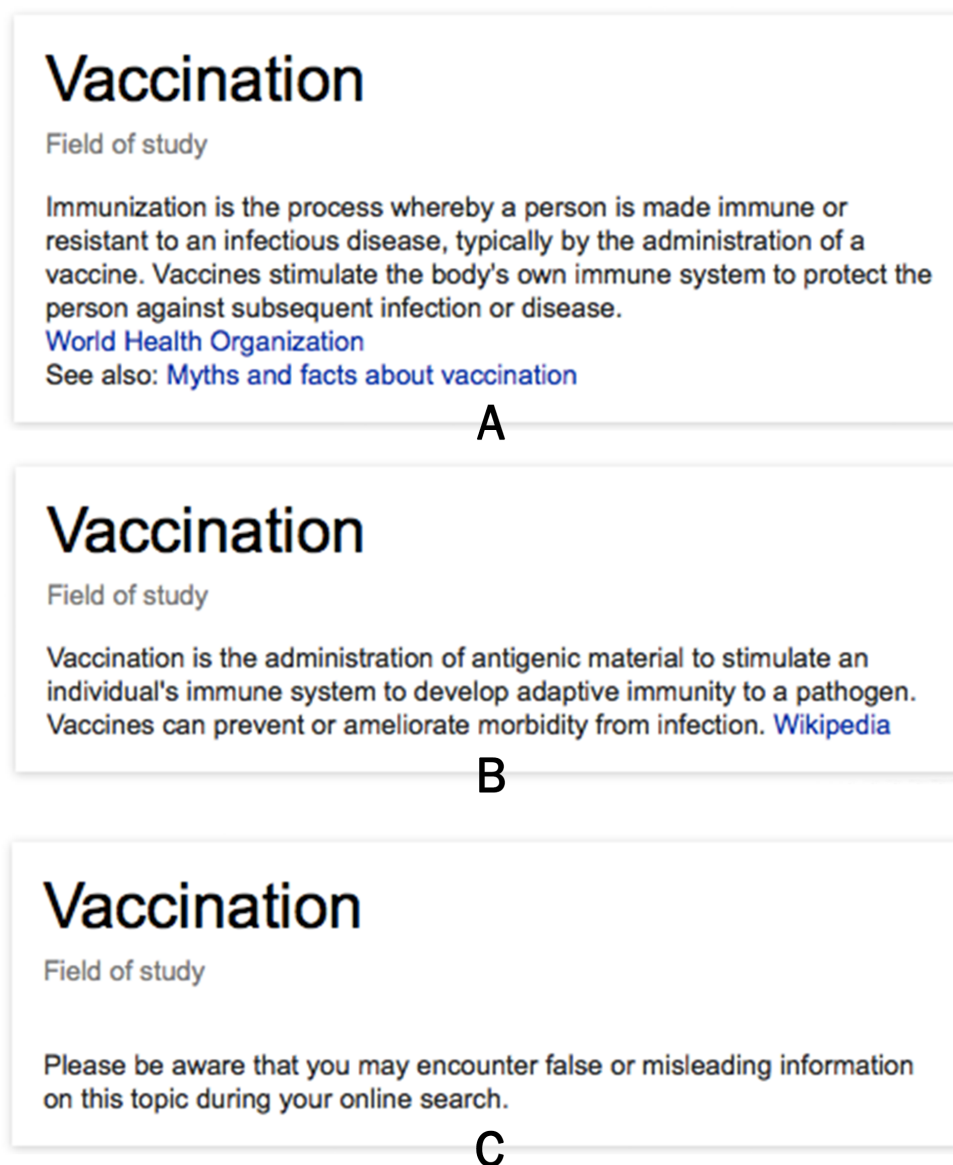
Search Engine Manipulation

In all six groups, Google was manipulated to display a ratio of 50% pro- and 50% antivaccination websites as search results in a randomized order. This manipulation was realized by configuring the context and annotation files of the Google custom search engine [23,51]. In short, two custom search engines were built: one restricted to search from a set of pro-vaccination websites (ProVaccineSearcher) and the other to a set of antivaccination websites (ConVaccineSearcher). Thereby, the set of websites belonging to the ProVaccineSearcher contained only websites that had an overall positive standpoint toward vaccination, while the ConVaccineSearcher consisted of a set of webpages with an overall negative standpoint toward vaccination. All websites were content analyzed and subsequently categorized following the descriptions in [23,52]. By using JavaScript and the Google custom search application programming interface (API) [51], we controlled the search execution of both search engines and

the display of the search results delivered from both [53]. Once a user entered a search query that matched the topic of the search task, the two engines were launched with the same query. The retrieved results were the first five relevant items (as chosen by the unmanipulated Google retrieval algorithm) from the pro-vaccination Web domains and similarly the first five items from the antivaccination Web domains. The retrieved results from both sources were combined and shuffled in a random order before being displayed. This was done for the construction of every search results page. Overall, there were 10 search results pages, with each page displaying 10 results, which mirrors Google's default search template. The random shuffling guaranteed control of any order effects in the display of the pro- and antivaccination search results.

Recruitment and Participants

A Web platform using the Drupal framework, which is an open source content management system [54], was developed and customized for the experiment. We designed a human

intelligence task (HIT) asking potential participants from CrowdFlower [55] to go to the Web platform we prepared. Prior to the actual recruitment, we pretested the experiment in two consecutive rounds, again using CrowdFlower, to check for the functioning of the manipulation's components, usability, and task understanding. After this pretest, we proceeded with the launching of the data collection. The recruitment advanced in sequential batches of on average 50 participants per HIT and lasted for 3 weeks in early 2015. Participants, registered as CrowdFlower workers, who passed the qualification requirements described subsequently were able to preview the HIT and apply for it. On completion of each batch, we analyzed the submissions and accepted people who followed the experiment as explained. There were two qualification requirements for the HIT: (1) users had to have the highest quality rating (Level 3, that is users who showed to be reliable and demonstrated high performance in completing posted tasks on the CrowdFlower platform), and (2) users needed to be located in an English-speaking country. The duration for task completion was set to 30 minutes.

Experimental Procedure

Once the HIT was published on the CrowdFlower platform, qualified users were able to select it. The HIT forwarded the user to the Web platform that proceeded in several steps in multiple sliding screens. In the first screen, a short introduction to the experiment was presented. After giving informed consent and the CrowdFlower ID on the next page, participants were presented with a screen explaining the search task. Participants were randomly assigned to one of the five experimental or the control groups. All participants were asked to search for information about vaccination for 10 minutes and started their search on the manipulated version of Google. They were free to enter any kind of vaccination-related keyword that came to their mind. Examples of entered search queries are "vaccination," "measles vaccine," "flu vaccine," or "vaccination pros and cons." A word cloud displaying the entered search queries according to their frequency is displayed in Multimedia Appendix 2. If participants entered keywords that were not related to vaccination, they received a message reminding them of the relevant search topic. This was done in two ways: (1) by pre-processing the entered search queries and matching the entered keywords to a regular expression that included the various inflected forms of the vaccination concept including the misspelled ones and/or by (2) extracting the noun phrases and using the comprising keywords to query the freebase API [56] for determining the relevance of the search query to viral and infectious diseases. This allowed for matching vaccine names (eg, "mmr," "measles," or "chickenpox"). During the search, participants could enter as many search queries as they wished and open all search results they were interested in. Participants in the experimental groups received, in addition to the search results, a knowledge graph box that was displayed on the upper right-hand side of the screen. After 10 minutes, participants were guided to the questionnaire that took around 10-15 minutes to complete. Participants were debriefed through the CrowdFlower platform after the data collection had ended.

The platform was designed to capture users' search behavior by recording their issued search queries, mouse hovers, clicks and selection of search elements, and their transition from one screen to another during the whole experiment. Video demos of the experimental workflow and manipulation can be found in Multimedia Appendix 3.

Measures

Posttest Items

Participants were asked to fill in a posttest-questionnaire after the search. The questionnaire was pretested to avoid biased results due to language issues and contained the following measures used in the subsequent analyses:

1. Perceived utility of the knowledge graph box: semantic differential consisting of 6 items such as helpful/disturbing or comprehensible/ambiguous.

2. Perceived quality of the information in the knowledge graph box: Battery of 7 items asking, for example, whether the information in the knowledge graph box was considered relevant (0=highly irrelevant to 6=highly relevant), credible (0=I completely agree to 6=I completely disagree), or comprehensible (0=I completely agree to 6=I completely disagree), adapted from [23].

3. Evaluation of the websites: Battery of 10 items asking, for example, whether the information found on the websites was considered relevant (0=highly irrelevant to 6=highly relevant), credible (0=I completely agree to 6=I completely disagree), or comprehensible (0=I completely agree to 6=I completely disagree), adapted from [23].

4. Beliefs and attitudes toward vaccination: Set of 7 items consisting of statements about risks and benefits of vaccination that were answered on a Likert scale (1=completely disagree to 7=completely agree). Two further items measured the perceived likelihood of severe side effects after a child or an adult got vaccinated (0=very unlikely to 6=very likely), all adapted from [23].

5. Knowledge about vaccination: Validated scale consisting of 9 true/false items [57]. Knowledge scores were calculated for each participant.

6. Sociodemographic information: 9 items asking participants about, for example, their gender, age, educational level, nationality, vaccination status, profession, or perceived confidence in their information seeking skills.

Measures 1 and 2 were applied only in the experimental groups. The complete questionnaire is provided in Multimedia Appendix 4.

Outcome Variables

Knowledge About Vaccination

A knowledge score was computed for every participant by summing the scores of all items of the knowledge about vaccination measure [57].

Beliefs and Attitudes Toward Vaccination

An exploratory factor analysis was run on the posttest items measuring participants' beliefs and attitudes toward vaccination and their evaluation of the search results with the goal to identify

the latent variables, respective constructs. Given that some of the items were added or slightly modified from the ones used in [23], we aimed at finding the latent constructs using the responses from the current experiment and consequently obtaining factor scores using the items loaded on the emerging factors. Additionally, the emerging factors were validated by comparing them and their corresponding item loadings to the ones reported in [23].

Evaluation of the Knowledge Graph Box

Perceived Utility of the Knowledge Graph Box

The perceived utility of the knowledge graph box was assessed by computing a total sum score of the posttest items reported in Section 1. The item responses were coded from 0-6 and hence, the perceived utility score ranged from 0-36. The higher the score, the higher the perceived utility of the knowledge graph box.

Perceived Quality of the Information in the Knowledge Graph Box

The second measure was represented by the sum of the posttest items in Section 2, each measuring trust, correctness, persuasiveness, relevance, credibility, and comprehensibility of the information embedded in the knowledge graph box. Similar to the first evaluation measure (perceived utility), the responses were coded from 0-6. This measure represented the perceived quality of the knowledge graph box's information with a total score ranging from 0-36. Again, a higher score indicated a higher perceived quality of the information.

Data Analysis

Randomization check was completed to investigate if there were differences among the groups that should be taken into account in subsequent analyses. Results were insignificant for all items measuring sociodemographic variables, implying that there were no significant differences among the experimental groups (see Multimedia Appendix 5).

To test Hypotheses 1-4 and answer Research Question 1, two-way analyses of variance (ANOVAs) were conducted, one for each outcome variable with availability of basic information and the presence of a warning as independent variables.

Research Question 2 was answered by using two-way ANOVAs for the dependent variables perceived utility of the knowledge graph box and perceived quality of the information embedded in the knowledge graph box with the availability of basic information and the presence of a warning as independent variables.

Two-way ANOVAs were chosen as the procedure for analysis since we aimed at investigating the separate main effects as well as the interaction effect of both independent variables on the dependent variables.

Statistical analyses were considered significant at $P < .05$. Moreover, post-hoc analysis was conducted for the observed significant results, if applicable. Bonferroni correction and Tukey HSD were applied to control for the family-wise error such as increased probability of Type I error due to multiple comparisons. The analysis of self-reported measures was conducted using SPSS 20.

Additionally, we looked at heatmaps that we generated from the recorded search hovers to detect if hovering patterns over the knowledge graph box are different among the experimental groups.

Results

The sample size was N=279 (group 1: n=45, group 2: n=45, group 3: n=46, group 4: n=50, group 5: n=47, group 6 [control group]: n=46). Over half (54.8%, 153/279) of participants were male and 45.2% (126/279) were female. The majority of participants came from the United States (43.4%, 121/279), followed by 29.4% (82/279) from the United Kingdom, 18.6% (52/279) from Canada, 1.8% (5/279) from Australia, and 1.1% (3/279) from New Zealand. The remaining respondents came from other countries. Just over half (52.0%, 145/279) of the participants were college graduates or had completed postgraduate studies, 22.2% (62/279) had concluded some college level, 5.7% (16/279) had received post-high school vocational or technical training, and 20.1% (56/279) had at least 8 years of high school or a completed high school degree. The mean age of participants was 37.34 years (SD 10.67), with 19 years as minimum and 69 years as maximum age.

Knowledge About Vaccination

Using two-way ANOVA, there was a significant main effect of availability of basic information in the knowledge graph box on participants' knowledge ($F_{2,273}$=4.86, P=.01). Presence of the warning did not have a significant main effect on knowledge, nor was there an interaction effect. Following up with a post-hoc analysis with Bonferroni correction and Tukey HSD, participants belonging to Groups 1 and 4 receiving comprehensible information with or without the warning, combined, had significantly higher mean scores on the knowledge scale compared to the participants in Groups 2 and 5 who were exposed to the hardly comprehensible information (mean difference 1.18, P=.01). Figure 3 shows the estimated marginal means of the knowledge score by both experimental factors. It is worth noting that the warning increased knowledge in participants who received the comprehensible information while it decreased knowledge in those who received the hardly comprehensible information as well as in those who did not get information at all. Hence, Hypothesis 1 is partially supported while Hypothesis 3 is not.

Figure 3. Posttest knowledge about vaccination as dependent on presence of a warning and comprehensibility level of basic information.
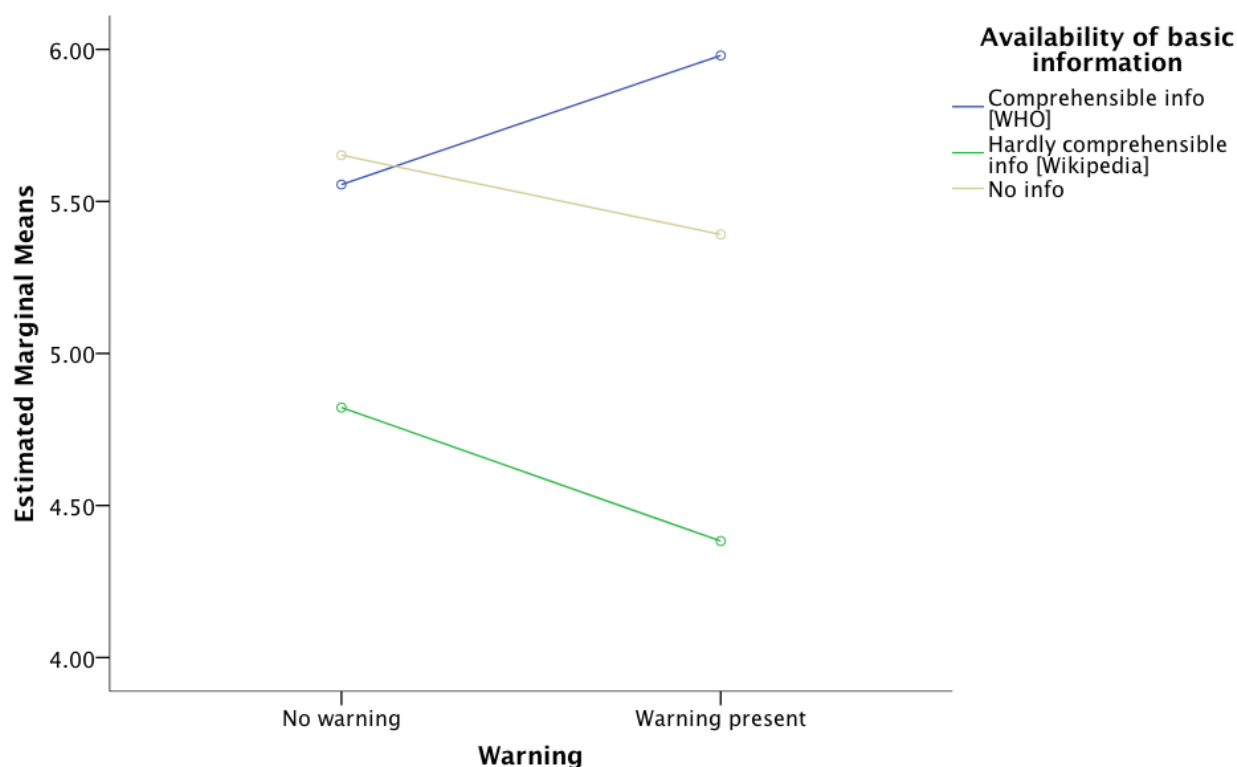
Beliefs and Attitudes Toward Vaccination

Factor Analysis

Factor analysis was run on 19 items (posttest items in Sections 3 and 4) with oblique rotation (promax) using the maximum likelihood extraction method. The Kaiser-Meyer-Olkin (KMO) measure verified the analysis's sampling adequacy (KMO=.90). Bartlett's test of sphericity (χ^2_{153}=3495.26, *P*=.001) indicated that correlations between items were sufficiently large for factor analysis. An initial analysis was run to obtain eigenvalues for each component in the data. Three eigenvalues were greater than Kaiser's criterion of 1 and in combination explained 65% of the variance. Based on the parallel analysis and the scree plot in which an inflection point was detected, three factors were retained for final analysis. To assess the reliability of the subscales that emerged from factor analysis, we measured Cronbach's alpha. Table 2 shows the pattern matrix of the factor analysis including each item's loadings on the obtained three factors. The eigenvalues were equal to 7.3, 3.9, and 1.1, and the percentage of variance explained by each was 38.4%, 20.7%, and 5.8%. In addition, Cronbach's alpha for each subscale comprising the items that showed loadings greater than or equal to 0.4 was .91, .89, and .89, respectively. The items forming factor 1 suggest a representation of skepticism or fear of vaccination side effects. Factor 2 displays an evaluation of information quality and factor 3 the acknowledgment of vaccination benefits. Factors 1 and 3 were negatively correlated (-.71), while both showed an insignificant positive correlation (.13, .05) with factor 2. The emerged factors and comprised items mirror the ones reported in [23].

Table 2. Exploratory factor analysis of beliefs and attitudes measures.

Pattern matrix[a]	Factor loadings		
	1	2	3
In your opinion, how likely is the occurrence of serious side effects after a child got vaccinated?	0.85[b]		
When I read about the possible side effects of vaccination on the websites, I felt worried.	0.82[b]		
When recommending vaccination, physicians do not pay enough attention to the possible side effects.	0.79[b]		
In your opinion, how likely is the occurrence of serious side effects after an adult got vaccinated?	0.77[b]		
When I read what the websites said about the effectiveness of vaccination, I felt worried.	0.76[b]		
Many vaccinations today do more harm than good.	0.75[b]		
Many of the vaccinations recommended today are redundant because the disease is almost extinct.	0.64[b]		
Vaccination often does not fully protect against a disease.	0.52[b]		
Thinking about the 10 minutes you spent searching for information on vaccination: Has the search made you more skeptical about vaccination or has the experience increased your confidence in it?	-0.38		
How convincing did you find the websites you looked at during your search?		0.91[b]	
The information on the websites I read was credible.		0.88[b]	
How much do you trust the information about vaccination you found on the websites before?		0.86[b]	
Was the information you found on the websites relevant?		0.73[b]	
How much do you trust Google to provide you with good information?		0.63[b]	
The information about vaccination I have just read on the websites was comprehensible for me.		0.52[b]	
If it weren't for vaccination, many people would have a shorter lifespan today than they do.			0.86[b]
People who opt out of vaccination do not only put themselves at risk, but also other people.			0.74[b]
Vaccination is one of the greatest medical breakthroughs affecting our lives.			0.74[b]
In my opinion, people should follow the advice to get vaccinated.			0.60[b]

[a]Extraction method: maximum likelihood. Rotation method: Promax with Kaiser normalization.

[b]Loading ~ 0.4.

Factor 1: Skepticism/Fear of Vaccination Side Effects

Two-way ANOVA was used for investigating the effect of the two experimental factors (availability of basic information and warning message) on the scores of the first factor (skepticism/fear of vaccination side effects). The only detected significant effect was the main effect of the availability of basic information on the factor 1 scores ($F_{2,273}$=3.5, P=.03).

Following up with a post-hoc analysis using Tukey HSD, participants receiving a knowledge graph box displaying comprehensible information (Groups 1 and 4) had significantly lower mean scores compared to the ones getting a knowledge graph box containing hardly comprehensible information

(Groups 2 and 5; mean difference -0.36, P=.047). However, when using Bonferroni correction, the P value was equal to .05 at the threshold of rejecting the null hypothesis. This result is a consequence of the conservative nature of Bonferroni correction, which is used to control for Type I error by testing on a lower alpha-level that is, in our case, equal to 0.016 (eg, 0.05/3). Overall, participants exposed to comprehensible information with or without the warning were less afraid/skeptical of vaccination side effects compared to the ones receiving less comprehensible information, again regardless of the presence of a warning message. Figure 4 shows that the warning message reduces the skepticism among the group exposed to the comprehensible information. However, the opposite is true for the other groups.

Figure 4. Skepticism/fear of vaccination side effects as dependent on presence of a warning and comprehensibility level of basic information.
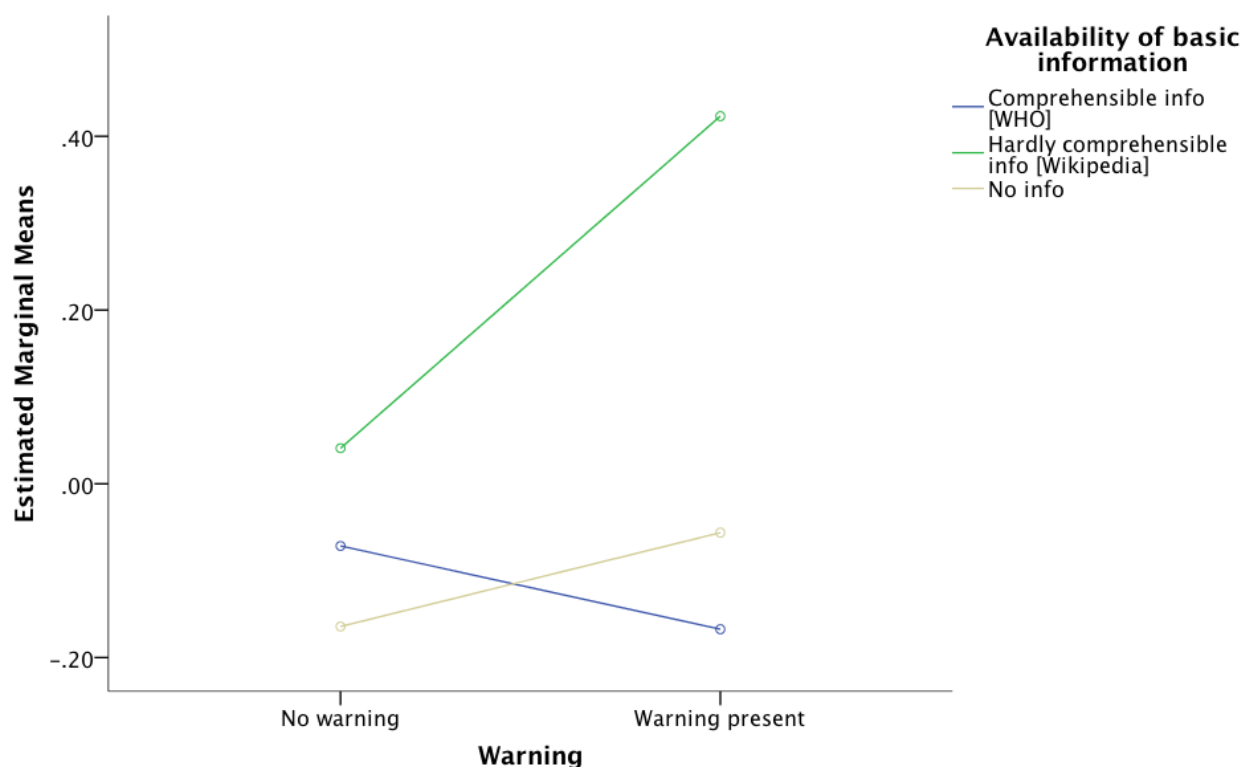
### Factor 2: Information Quality

Using two-way ANOVA, the only detected significant effect was the main effect of availability of basic information on the scores of factor 2 ($F_{2,273}$=3.73, P=.02). Following up with a post-hoc analysis, participants belonging to the groups receiving a knowledge graph box with comprehensible information (Groups 1 and 4) had significantly lower mean scores as compared to both groups getting a knowledge graph box displaying hardly comprehensible information (Groups 2 and 5; mean difference -0.36, P=.049 using Bonferroni correction

and P=.043 using Tukey HSD). In other words, the former groups seem to be more critical in terms of the evaluation of information quality compared to the latter groups. Figure 5 shows the comparison between groups receiving the same type of basic information and the ones that additionally saw a warning message. Again, the warning appears to have a positive influence on the group receiving comprehensible information. The reverse pattern emerges for the group being exposed to hardly comprehensible information and slightly also for the group not getting basic information.

Figure 5. Attitude toward information quality as dependent on presence of a warning and comprehensibility level of basic information.
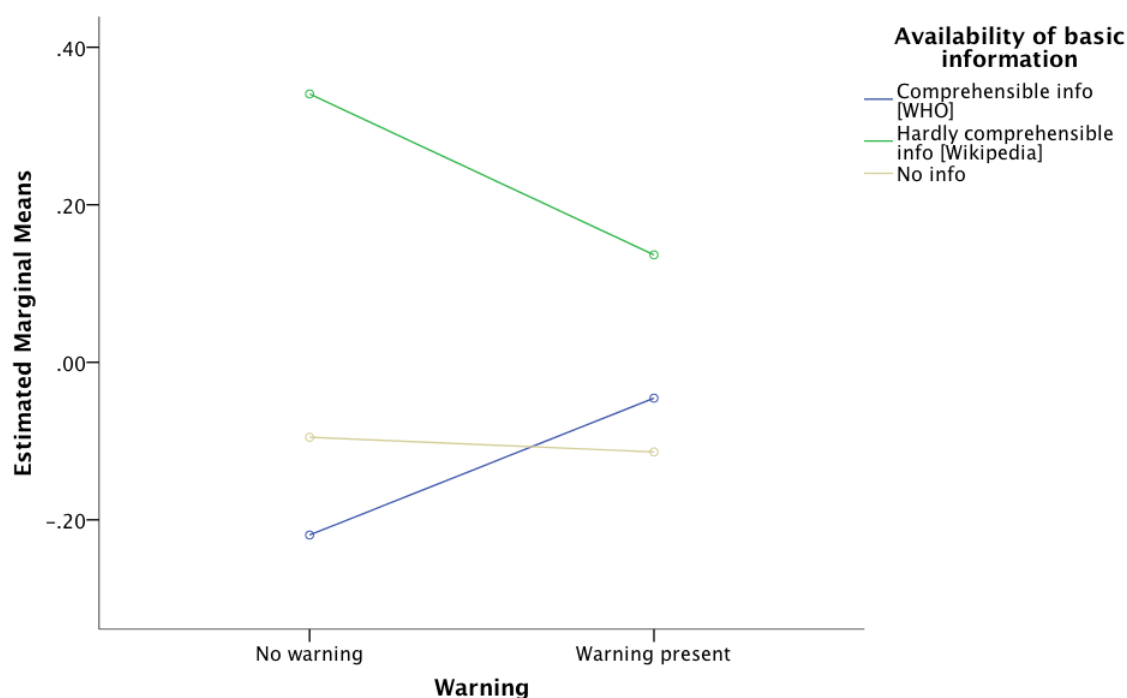
Factor 3: Acknowledgment of vaccination benefits

There was no significant effect with respect to the main effects of either experimental factor or for the interactions on the scores of Factor 3. However, groups receiving comprehensible basic information in the knowledge graph box, regardless of the presence of a warning, had the highest average scores on Factor 3 indicating a higher appreciation of vaccination benefits (see Figure 6).

Thus, Hypothesis 2 is partially supported while Hypothesis 4 is not. Moreover, there was no significant interaction (on all levels of experimental variables) between the availability of information and the warning on the outcome measures (RQ1). However, the overall results suggest that there was an interaction with regard to the groups receiving basic information (comprehensible and hardly comprehensible) and the corresponding groups with the added warning message.

Figure 6. Acknowledgment of vaccination benefits as dependent on presence of a warning and comprehensibility level of basic information.
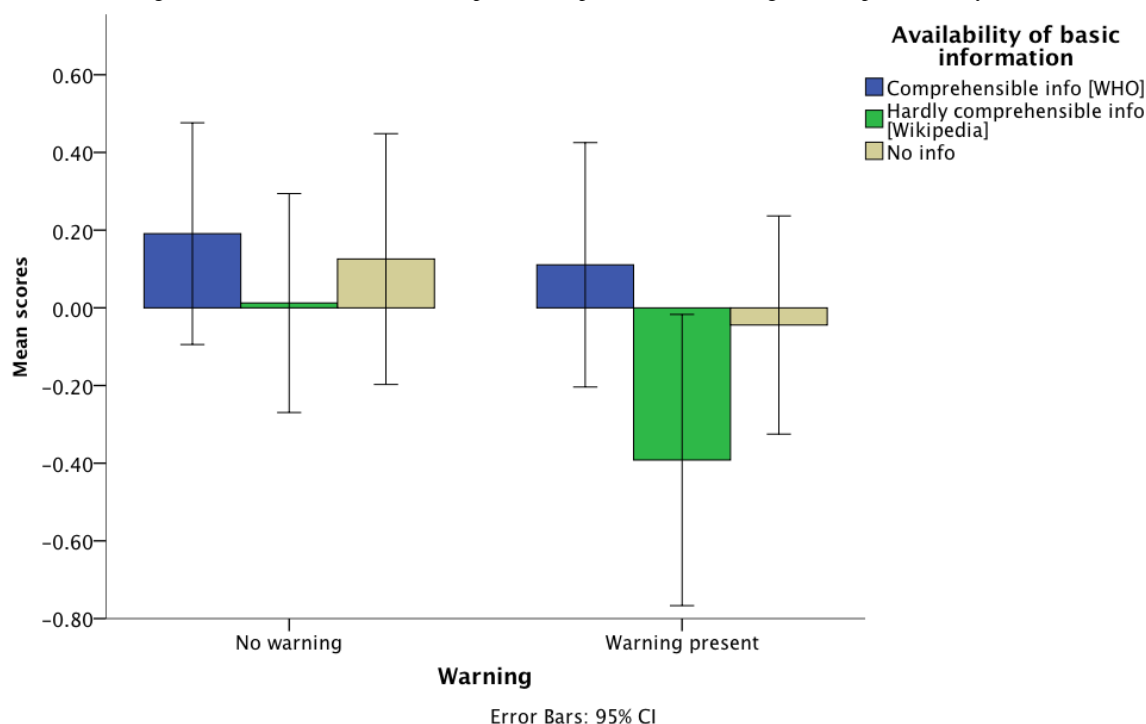
Evaluation of the Knowledge Graph Box

Two-way ANOVA was conducted for each dependent variable (perceived utility and perceived quality of the information displayed by the knowledge graph box) with the availability of basic information and the presence of a warning as independent variables. This time, five groups were compared since the control group was not exposed to the knowledge graph box and hence did not answer the items evaluating it.

In both cases, there was no significant effect with respect to the main effects of either experimental factor or of the interactions on the scores of the dependent variables. On average, the groups scored 24.4 out of 36 on perceived utility and 23.3 out of 36 on perceived information quality embedded in the knowledge graph box. This amounts to an average evaluation of 67.7% and 64.7%, suggesting an overall moderate to positive evaluation of the knowledge graph box. Heatmaps generated from the aggregated mouse hovers per group over the search elements showed similar hovering patterns among the groups (see Multimedia Appendix 6).

Discussion

Principal Findings

The findings suggest a positive effect of comprehensible basic information on information seekers' vaccination-related knowledge and attitudes and a moderate to positive evaluation of the knowledge graph box among all groups. In contrast, the sole presentation of a warning message did not result in any substantial effects. Also, there was no significant interaction between the basic factual information and the warning.

The observation of a positive influence of easily comprehensible text is in line with previous research suggesting a higher persuasiveness of comprehensible information [44,45]. In this study, comprehensibility might have had a persuasive character in the sense of fostering participants' agreement with considering vaccination as a form of health protection. The warning message, however, did not yield comparable effects. This could be explained by a lacking reference point for the information seekers. Literature on warning messages came to the conclusion that a warning is most successful when it is combined with advice for action such as a strategy to avoid the threat [48]. Thus, the warning message by itself might have disrupted consumers' biased search but since a strategy stating how to handle or detect misleading information was missing, they could not translate this knowledge about the presence of false information into more successful search outcomes.

Interestingly, the combination of comprehensible information and a warning message supports this assumption, although on a non-significant level. That is, the obtained results suggest that comprehensible information combined with a warning steers consumers toward a vaccination-supporting position. Since the comprehensible information might have been interpreted by the participants as promoting vaccination, the alert could have been linked to the information and consequently been perceived as warning against antivaccination information. In contrast, the combination of hardly comprehensible information and a warning seems to direct information seekers more toward a

vaccination-skeptical position. This result might be, on the one hand, due to the content of the Wikipedia article that was linked to the knowledge graph box. When clicking on it, participants did not only see information about the importance of vaccination and factual information (eg, history, functioning mechanisms, statistics) but also paragraphs related to side effects, adverse events, and the surrounding controversy. This might have sparked concerns about vaccination. On the other hand, for laypersons, the hardly understandable definition might have evoked confusion and insecurity among the health information seekers. In combination with a warning lacking polarity or a clear positioning, these doubts might have been augmented, leading to the opposite effect of the warning as compared to the group receiving comprehensible basic information. A follow-up on the true relationship between these two experimental factors would be an interesting starting point for future research in this area.

The comparison of the results to the ones of a recent study investigating the effects of different ratios of pro- and antivaccination information displayed by Google [23] stresses the potential of this intervention in terms of effectiveness and applicability. When looking at the achieved factor scores, the groups exposed to comprehensible information in this experiment achieved better outcomes than the groups being exposed to an almost equal amount of pro- and antivaccination information in [23] as they were more knowledgeable, less skeptical of vaccination, and more critical of information quality. This is promising because it shows that filtering the retrieved search results according to their quality and source credibility is not the only approach to minimize the negative effects of a biased information search. Indeed, also the incorporation of comprehensible semantic information of the health-related search topic in the knowledge graph box appears to be successful, while being more feasible.

Limitations

Although this study offers valuable insights into a new technological debiasing technique, some limitations need to be considered. First, we employed a design that immediately tested the effectiveness of a knowledge graph box as a debiasing technique without a prior phase that only observed the occurrence of a biased search and information processing. Acknowledging the solid body of research that demonstrates the deficiency of people's information search behavior and processing [23,24,28,33-35], we took the problem context for granted, rendering another demonstration of the same phenomenon unnecessary.

A second limitation refers to the recruitment of participants through CrowdFlower. Although the sample was sufficiently diverse concerning sociodemographic characteristics such as age, gender, and nationality, a disproportionate share of participants was highly educated. However, Internet samples are overall more heterogeneous than traditional samples [58]. Yet, the probably higher level of online search experience and computer skills remains a limitation to the generalization of the findings. Using an online sample further bears the peril of "non-serious" or "repeat responders" [58]. Still, a biasing effect of such participants on the results of the study was minimized

through a careful investigation and follow-up of each respondent. This was done in two ways. First, the developed platform captured participants' interaction and behavior throughout the course of the experiment. This information was then used to construct a timeline for every participant to verify that the experiment was conducted in one sitting. Second, the respondents were identified during the experiment based on their inserted CrowdFlower ID. This allowed for a validation of the recorded data for every participant in our platform against the task-submission data on the CrowdFlower platform.

Moreover, the template Google currently uses when displaying the information in the knowledge graph box includes both the basic information and its source as a hyperlink. As a result, we cannot establish or compare a difference in contribution of the source and the information itself to the observed effect. Our emphasis on the importance of the information itself is, however, in line with the communication science-based assumption that the content of a message is more crucial than variables such as source credibility or attractiveness [43]. This assumption is also empirically supported by the equally high evaluation scores awarded to the knowledge graph box by the experimental groups, regardless of the displayed source. Nevertheless, this was only a first test of the possible effects of the knowledge graph box focusing on the differentiation between its overall presence and two content-related factors, namely comprehensibility of basic information and presence of a warning. Future research could experimentally distinguish between single components of the knowledge graph box and investigate a potential influence of the content's source more thoroughly.

Eventually, we could have employed a pre-posttest design instead of a posttest-only design. Using a different experimental design would have allowed for assessing participants' vaccination-related baseline knowledge and attitudes. However, we refrained from this alternative in order to not sensitize our participants to the experiment's true outcome of interest, which could have had a biasing effect. Further, we aimed at comparing the findings to the studies reported in [23] by keeping the experimental design constant. Still, further research investigating the online health information seeking process by applying different experimental designs is needed.

Conclusions

This study assessed the debiasing effects of a knowledge graph box providing health information seekers with basic information (comprehensible vs hardly comprehensible) and/or a warning message during their search on vaccination via Google. We intended to ensure a realistic setting of the experiment so that it could be implemented in real-life. That means we adhered to the technology currently used by Google to retrieve the semantic information for the knowledge graph box and used neutral information to design the experimental stimuli.

Since Google already uses the knowledge graph box [42], a simple change in the retrieval procedure for its content could make a valuable difference for the consumers. In fact, while this experiment was in the field, Google announced an update regarding the information displayed in the knowledge graph box when it comes to searching for health-related information [59]. The update concerns the retrieval of semantic information from "high-quality" websites such as the Mayo Clinic, WebMD, or Centers for Disease Control and Prevention [59]. However, this change pertained only to the United States and searches conducted in English [59]. Surprisingly, the update did not affect searches conducted on vaccination- or vaccine-related search terms at the time of writing this manuscript (see Multimedia Appendix 1).

Hence, the current situation suggests that there is still an extensive potential to improve the online health information seeking process, as well as a necessity for further research in this area. Instead of developing and studying complex interventions to improve people's health information search, we suggest the further investigation of more realizable strategies like the one presented here.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Video demo of actual search on Google about vaccination (February 2016).

[MP4 File (MP4 Video), 13MB - jmir_v18i6e137_app1.mp4]

Multimedia Appendix 2

Word cloud of entered search queries.

[PNG File, 376KB - jmir_v18i6e137_app2.png]

XSL•FO

RenderX

Multimedia Appendix 3

Video demo of experimental groups.

[MP4 File (MP4 Video), 15MB - jmir_v18i6e137_app3.mp4]

Multimedia Appendix 4

Complete questionnaire.

[PDF File (Adobe PDF File), 644KB - jmir_v18i6e137_app4.pdf]

Multimedia Appendix 5

Results of the randomization check.

[PDF File (Adobe PDF File), 51KB - jmir_v18i6e137_app5.pdf]

Multimedia Appendix 6

Heatmaps showing the aggregated mouse hovering.

[PDF File (Adobe PDF File), 2MB - jmir_v18i6e137_app6.pdf]

References

1. Lewandowsky S, Ecker UKH, Seifert CM, Schwarz N, Cook J. Misinformation and Its Correction: Continued Influence and Successful Debiasing. Psychol Sci Public Interest 2012 Dec;13(3):106-131. [doi: 10.1177/1529100612451018] [Medline: 26173286]
2. Fox S, Duggan M. Health Online 2013. Washington, DC: Pew Internet and American Life Project; 2013. URL: http://www.pewinternet.org/2013/01/15/health-online-2013/ [accessed 2015-11-17] [WebCite Cache ID 6d6YpwmtZ]
3. Flash Eurobarometer 404. European citizens' digital health literacy. A report to the European Commission. Brussels; 2014 Nov. URL: http://bookshop.europa.eu/en/european-citizens-digital-health-literacy-pbKK0414968/?CatalogCategoryID=CXoKABst5TsAAAAEjepEY4e5L [accessed 2015-11-17] [WebCite Cache ID 6d6ZzjjUY]
4. Cline RJ, Haynes KM. Consumer health information seeking on the Internet: the state of the art. Health Educ Res 2001 Dec;16(6):671-692 [FREE Full text] [Medline: 11780707]
5. Stern AM, Markel H. The history of vaccines and immunization: familiar patterns, new challenges. Health Aff (Millwood) 2005 May;24(3):611-621 [FREE Full text] [doi: 10.1377/hlthaff.24.3.611] [Medline: 15886151]
6. Andre FE. Vaccinology: past achievements, present roadblocks and future promises. Vaccine 2003 Jan;21(7-8):593-595. [doi: 10.1016/S0264-410X(02)00702-8]
7. Bonanni P. Demographic impact of vaccination: a review. Vaccine 1999 Oct;17:S120-S125. [doi: 10.1016/S0264-410X(99)00306-0]
8. Fenner F, Henderson DA, Arita I, Jezek Z, Ladnyi ID. Smallpox and its Eradication. Geneva: World Health Organization 1988 URL: http://apps.who.int/iris/handle/10665/39485 [accessed 2015-11-17] [WebCite Cache ID 6d6vKqQxH]
9. Kata A. A postmodern Pandora's box: anti-vaccination misinformation on the Internet. Vaccine 2010 Feb 17;28(7):1709-1716. [doi: 10.1016/j.vaccine.2009.12.022] [Medline: 20045099]
10. Davies P, Chapman S, Leask J. Antivaccination activists on the world wide web. Arch Dis Child 2002 Jul;87(1):22-25 [FREE Full text] [Medline: 12089115]
11. Zimmerman RK, Wolfe RM, Fox DE, Fox JR, Nowalk MP, Troy JA, et al. Vaccine criticism on the World Wide Web. J Med Internet Res 2005;7(2):e17 [FREE Full text] [doi: 10.2196/jmir.7.2.e17] [Medline: 15998608]
12. Wolfe RM, Sharp LK. Vaccination or immunization? The impact of search terms on the internet. J Health Commun 2005 Sep;10(6):537-551. [doi: 10.1080/10810730500228847] [Medline: 16203632]
13. Ward JK, Peretti-Watel P, Larson HJ, Raude J, Verger P. Vaccine-criticism on the internet: new insights based on French-speaking websites. Vaccine 2015 Feb 18;33(8):1063-1070. [doi: 10.1016/j.vaccine.2014.12.064] [Medline: 25583384]
14. Petts J, Niemeyer S. Health risk communication and amplification: learning from the MMR vaccination controversy. Health, Risk & Society 2004 Mar;6(1):7-23. [doi: 10.1080/13698570410001678284]
15. Gangarosa EJ, Galazka AM, Wolfe CR, Phillips LM, Miller E, Chen RT. Impact of anti-vaccine movements on pertussis control: the untold story. The Lancet 1998 Jan;351(9099):356-361. [doi: 10.1016/S0140-6736(97)04334-1]
16. Dubé E, Laberge C, Guay M, Bramadat P, Roy R, Bettinger JA. Vaccine hesitancy: An overview. Hum Vaccin Immunother 2013;9(8):1763-1773. [Medline: 23584253]
17. Smith PJ, Humiston SG, Marcuse EK, Zhao Z, Dorell CG, Howes C, et al. Parental delay or refusal of vaccine doses, childhood vaccination coverage at 24 months of age, and the Health Belief Model. Public Health Rep 2011 Aug;126 Suppl 2:135-146 [FREE Full text] [Medline: 21812176]

XSL•FO

RenderX

18. Gowda C, Dempsey AF. The rise (and fall?) of parental vaccine hesitancy. Hum Vaccin Immunother 2013 Aug;9(8):1755-1762 [FREE Full text] [doi: 10.4161/hv.25085] [Medline: 23744504]

19. Dubé E, Vivion M, MacDonald NE. Vaccine hesitancy, vaccine refusal and the anti-vaccine movement: influence, impact and implications. Expert Rev Vaccines 2015 Jan;14(1):99-117. [doi: 10.1586/14760584.2015.964212] [Medline: 25373435]

20. 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]

21. O'Keeffe J, Willinsky J, Maggio L. Public access and use of health research: an exploratory study of the National Institutes of Health (NIH) Public Access Policy using interviews and surveys of health personnel. J Med Internet Res 2011;13(4):e97 [FREE Full text] [doi: 10.2196/jmir.1827] [Medline: 22106169]

22. Wang L, Wang J, Wang M, Li Y, Liang Y, Xu D. Using Internet search engines to obtain medical information: a comparative study. J Med Internet Res 2012;14(3):e74 [FREE Full text] [doi: 10.2196/jmir.1943] [Medline: 22672889]

23. Allam A, Schulz PJ, Nakamoto K. The impact of search engine selection and sorting criteria on vaccination beliefs and attitudes: two experiments manipulating Google output. J Med Internet Res 2014;16(4):e100 [FREE Full text] [doi: 10.2196/jmir.2642] [Medline: 24694866]

24. Pan B, Hembrooke H, Joachims T, Lorigo L, Gay G, Granka L. In google we trust: Users' decisions on rank, position, and relevance. J Comput-Mediat Comm 2007;12(3):801-823. [doi: 10.1111/j.1083-6101.2007.00351.x]

25. Zhang J, Dimitroff A. The impact of webpage content characteristics on webpage visibility in search engine results (Part I). Information Processing & Management 2005 May;41(3):665-690. [doi: 10.1016/j.ipm.2003.12.001]

26. Lau AY, Coiera EW. A Bayesian model that predicts the impact of Web searching on decision making. J Am Soc Inf Sci 2006 May;57(7):873-880. [doi: 10.1002/asi.20377]

27. Lau AY, Coiera EW. How do clinicians search for and access biomedical literature to answer clinical questions? Amsterdam: IOS Press; 2007 Presented at: 12th World Congress on Health (Medical) Informatics; Building Sustainable Health Systems; 2007; Amsterdam.

28. Wirth W, Böcking T, Karnowski V, Von Pape T. Heuristic and systematic use of search engines. J Comput-Mediat Comm 2007;12(3):778-800. [doi: 10.1111/j.1083-6101.2007.00350.x]

29. Wilke A, Mata R. Cognitive bias. In: Ramachandran VS, editor. The Encyclopedia of Human Behavior. London, UK: Academic Press; 2012:531-535.

30. Gigerenzer G. Fast and frugal heuristics: The tools of bounded rationality. In: Koehler DJ, Harvey N, editors. Blackwell Handbook of Judgment and Decision Making. Malden, UK: Blackwell Publishing Ltd; 2004:62-88.

31. Tversky A, Kahneman D. Judgment under Uncertainty: Heuristics and Biases. Science 1974 Sep 27;185(4157):1124-1131. [doi: 10.1126/science.185.4157.1124] [Medline: 17835457]

32. Gigerenzer G, Gaissmaier W. Heuristic decision making. Annu Rev Psychol 2011;62:451-482. [doi: 10.1146/annurev-psych-120709-145346] [Medline: 21126183]

33. Lau AY, Coiera EW. Do people experience cognitive biases while searching for information? J Am Med Inform Assoc 2007;14(5):599-608 [FREE Full text] [doi: 10.1197/jamia.M2411] [Medline: 17600097]

34. Keselman A, Browne AC, Kaufman DR. Consumer health information seeking as hypothesis testing. J Am Med Inform Assoc 2008;15(4):484-495 [FREE Full text] [doi: 10.1197/jamia.M2449] [Medline: 18436912]

35. Kayhan VO. Seeking health information on the web: positive hypothesis testing. Int J Med Inform 2013 Apr;82(4):268-275. [doi: 10.1016/j.ijmedinf.2012.12.004] [Medline: 23294509]

36. Fischhoff B. Debiasing. In: Kahneman D, Slovic P, Tversky A, editors. Judgment under uncertainty: Heuristics and biases. 24th ed. New York: Cambridge University Press; 1982:422-444.

37. Sanna LJ. Debiasing. In: Baumeister RF, Vohs KD, editors. Encyclopedia of Social Psychology. Thousand Oaks, CA: Sage; 2007:218-220.

38. Kaufmann L, Michel A, Carter CR. Debiasing Strategies in Supply Management Decision-Making. J Bus Logistics 2009;30(1):85-106. [doi: 10.1002/j.2158-1592.2009.tb00100.x]

39. Larrick R. Debiasing. In: Koehler DJ, Harvey N, editors. Blackwell Handbook of Judgment and Decision Making. Malden, UK: Blackwell Publishing Ltd; 2004:316-337.

40. Aczel B, Bago B, Szollosi A, Foldes A, Lukacs B. Is it time for studying real-life debiasing? Evaluation of the effectiveness of an analogical intervention technique. Front Psychol 2015;6:1120 [FREE Full text] [doi: 10.3389/fpsyg.2015.01120] [Medline: 26300816]

41. Lau AY, Coiera EW. Can cognitive biases during consumer health information searches be reduced to improve decision making? J Am Med Inform Assoc 2009;16(1):54-65 [FREE Full text] [doi: 10.1197/jamia.M2557] [Medline: 18952948]

42. Singhal A. Google Official Blog. 2012 May 16. Introducing the Knowledge Graph: things, not strings URL: https://googleblog.blogspot.ch/2012/05/introducing-knowledge-graph-things-not.html [accessed 2015-11-17] [WebCite Cache ID 6d6qRTPaZ]

43. Fishbein M, Ajzen I. Acceptance, yielding, and impact: Cognitive processes in persuasion. In: Petty RE, Ostrom TM, Brock TC, editors. Cognitive responses in persuasion. Hillsdale, NJ: Lawrence Earlbaum; 1981:339-359.

XSL•FO
RenderX

44.   Scharrer L, Bromme R, Britt MA, Stadtler M. The seduction of easiness: How science depictions influence laypeople's reliance on their own evaluation of scientific information. Learning and Instruction 2012 Jun;22(3):231-243. [doi: 10.1016/j.learninstruc.2011.11.004]

45.   Scharrer L, Britt MA, Stadtler M, Bromme R. Easy to Understand but Difficult to Decide: Information Comprehensibility and Controversiality Affect Laypeople's Science-Based Decisions. Discourse Processes 2013 Aug;50(6):361-387. [doi: 10.1080/0163853X.2013.813835]

46.   Flesch R. A new readability yardstick. J Appli Psych 1948;32(3):221-233. [doi: 10.1037/h0057532]

47.   Added Bytes. Readability-Score.com. 2016. URL: https://readability-score.com/ [accessed 2016-02-02] [WebCite Cache ID 6ezzXtgZ7]

48.   Xiao B, Benbasat I. Designing Warning Messages for Detecting Biased Online Product Recommendations: An Empirical Investigation. Inf Syst Res 2016 Feb;12(01):793-811. [doi: 10.1287/isre.2015.0592]

49.   Huang H, Kuo M. Looks familiar, appears more valid? The moderating role of computer-supported warnings between information repetition and decision outcome. Behaviour & Information Technology 2015 Jul 24;34(11):1119-1128. [doi: 10.1080/0144929X.2015.1069397]

50.   Wikidata. Welcome to Wikidata. 2015. URL: https://www.wikidata.org/wiki/Wikidata:Main_Page [accessed 2016-01-22] [WebCite Cache ID 6ejLPBRo9]

51.   Google Developers. Google Custom Search API. 2015. URL: https://developers.google.com/custom-search/ [accessed 2015-11-19] [WebCite Cache ID 6d9kq7Oym]

52.   Sak G, Diviani N, Allam A, Schulz PJ. Comparing the quality of pro- and anti-vaccination online information: a content analysis of vaccination-related webpages. BMC Public Health 2016;16(1):38 [FREE Full text] [doi: 10.1186/s12889-016-2722-9] [Medline: 26769342]

53.   Google Developers. Google custom search rendering. 2015. URL: https://developers.google.com/custom-search/docs/js/rendering [accessed 2015-11-19] [WebCite Cache ID 6d9l4gC2d]

54.   Drupal - Open Source CMS. Drupal Content Management Framework. 2016. URL: https://www.drupal.org [accessed 2016-02-15] [WebCite Cache ID 6fJl2rYwn]

55.   CrowdFlower crowdsourcing platform. 2015. URL: http://www.crowdflower.com/ [accessed 2015-11-19] [WebCite Cache ID 6d9lDC6Gl]

56.   Google Developers. Freebase API. 2015. URL: https://developers.google.com/freebase/?hl=en [accessed 2016-01-22] [WebCite Cache ID 6ejLSBI94]

57.   Zingg A, Siegrist M. Measuring people's knowledge about vaccination: developing a one-dimensional scale. Vaccine 2012 May 28;30(25):3771-3777. [doi: 10.1016/j.vaccine.2012.03.014] [Medline: 22445808]

58.   Gosling SD, Vazire S, Srivastava S, John OP. Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. Am Psychol 2004 Mar;59(2):93-104. [doi: 10.1037/0003-066X.59.2.93] [Medline: 14992636]

59.   Ramaswami P. A remedy for your health-related questions: health info in the Knowledge Graph. 2015. Google Official blog URL: https://googleblog.blogspot.ch/2015/02/health-info-knowledge-graph.html [accessed 2016-02-16] [WebCite Cache ID 6fL15iiuh]

Abbreviations

ANOVA: analysis of variance
API: application programming interface
HIT: human intelligence task
KMO: Kaiser-Meyer-Olkin
WHO: World Health Organization

XSL•FO
RenderX

Original Paper

Consumer Decision-Making Based on Review Websites: Are There Differences Between Choosing a Hotel and Choosing a Physician?

Fabia Rothenfluh[1], MSc; Evi Germeni[1], Ph.D.; Peter J Schulz[1], Ph.D.

Institute of Communication and Health, Faculty of Communication Sciences, Università della Svizzera italiana, Lugano, Switzerland

Corresponding Author:
Fabia Rothenfluh, MSc
Institute of Communication and Health
Faculty of Communication Sciences
Università della Svizzera italiana
Via Giuseppe Buffi 13
Lugano, 6900
Switzerland
Phone: 0041 58 666 4485
Fax: 0041 58 666 4647
Email: fabia.rothenfluh@usi.ch

Abstract

Background: Web users are increasingly encouraged to rate and review consumer services (eg, hotels, restaurants) and, more recently, this is also the case for physicians and medical services. The resemblance in the setup and design of commercial rating websites (CRWs) and Web-based physician rating websites (PRWs) raises the question of whether choice-making processes based on the two types of websites could also be similar.

Objective: This qualitative study sought to explore the extent to which consumer decision making based on Web-based reviews is the same for consumer services (ie, choice of a hotel) and health services (ie, choice of a pediatrician), while providing an in-depth understanding of potential differences or similarities.

Methods: Between June and August 2015, we carried out a total of 22 qualitative interviews with young parents residing in the German-speaking part of Switzerland. Participants were invited to complete 2 choice tasks, which involved (1) choosing a hotel based on the commercial Web-based rating website TripAdvisor and (2) selecting a pediatrician based on the PRW Jameda. To better understand consumers' thought processes, we instructed participants to "think aloud", namely to verbalize their thinking while sorting through information and reaching decisions. Using a semistructured interview guide, we subsequently posed open-ended questions to allow them to elaborate more on factors influencing their decision making, level of confidence in their final choice, and perceived differences and similarities in their search for a hotel and a physician. All interviews were recorded, transcribed, and analyzed using an inductive thematic approach.

Results: Participants spent on average 9:57 minutes (standard deviation=9:22, minimum=3:46, maximum=22:25) searching for a hotel and 6:17 minutes (standard deviation=4:47, minimum=00:38, maximum=19:25) searching for a pediatrician. Although the choice of a pediatrician was perceived as more important than the choice of a hotel, participants found choosing a physician much easier than selecting an appropriate accommodation. Four main themes emerged from the analysis of our interview data that can explain the differences in search time and choice confidence: (1) trial and error, (2) trust, (3) competence assessment, and (4) affect and likeability.

Conclusions: Our results suggest that, despite congruent website designs, individuals only trust review information to choose a hotel, but refuse to fully rely on it for selecting a physician. The design and content of Web-based PRWs need to be adjusted to better address the differing information needs of health consumers.

Introduction

Web-based review platforms allow consumers to post statements about products and services on the Internet by means of positive or negative reviews, also called electronic word of mouth [1]. Such websites are numerously accessible for purchases of physical goods and services (eg, amazon.com or zalando.de), as well as tourism and gastronomy (eg, booking.com or tripadvisor.com). Since the early 2000s, ratings have also been accessible for medical services, including the choice of physicians. Web-based physician rating websites (PRWs) "collect and present information about patients' experience and satisfaction with individual physicians and their practices" [2].

Previous research has shown that the use of PRWs has increased over the last decade. In 2012, about 36% of Americans reported to have searched for a physician on the Internet, whereas 65% were aware of such ratings [3]. In European countries, such as Germany, in contrast, far fewer people have heard of (32%) or used (25%) such rating websites [4]. In Switzerland, awareness and use are estimated to be even lower as patients can presently post only positive reviews about their physician because of legal restrictions, which are supposed to be lifted in 2016 or 2017 [5]. Yet, PRWs' impact on consumers' final consultation decisions seems to be considerable: 65% of German PRW users have consulted a physician based on the ratings provided by these websites [6], and 30% of American consumers have checked PRWs before consulting a physician [7]. Younger generations, in particular, seem to be increasingly relying on the Internet when selecting a doctor. More than a quarter of young parents surveyed in the United States indicated that they had selected the pediatrician for their child on the Internet [8,9].

Past research has pointed out ethical issues related to the use of PRWs [10-13,19] has focused on the prevalence and content of PRWs [2,6,7,11,14,16-18,20] and has provided an initial understanding of individuals' choice-making process when selecting a doctor on the Internet [8,15,43,44]. A study by Hanauer and colleagues [8] found a strong impact of physician ratings on patients' final choice. In a Web-based experiment, 22% of parents who were prompted to find a new pediatrician for their child had followed the advice of their neighbors, whereas 46% did so when the neighbors' recommendation was in line with positive online ratings of that physician. Nevertheless, only 3% consulted the pediatrician recommended by their neighbors if the doctor's Web-based reviews were negative [8]. Furthermore, Grabner-Kräuter and Waiguny [15] found that the review style (factual or emotional) and the review number influenced individuals' Web-based physician selection. Specifically, a high number of reviews resulted in a more positive attitude toward a physician, whereas the style mainly affected the reviewer's perceived expertise. These results have shed new light on the influence PRWs can have on patients' decision making. Yet, up to date, an in-depth understanding of how individuals choose a physician based on PRWs is missing. Similarly, the extent to which consumers' decision-making approach could be similar as for commercial rating websites (CRWs) is largely unknown.

The design and setup of CRWs holds significant similarities to that of PRWs (see Figure 1). Users can find general information, rate predefined aspects of products and services, write open-ended reviews, check the location and contact information, and see pictures. Due to the high visual similarity of PRWs and CRWs, we sought to investigate whether similarities could also exist with regard to the way that consumers reach their decisions based on these websites. Specifically, the objectives of the study were (1) to explore the extent to which the decision-making approach to selecting a physician is the same as for other consumer services or products (ie, choice of a hotel) and (2) to understand how potential differences in the search strategy can be explained.

Figure 1. Screenshots of the hotel rating website tripadvisor.de [21] (left) and the physician rating website jameda.de [22] (right) retrieved on March 2, 2016.
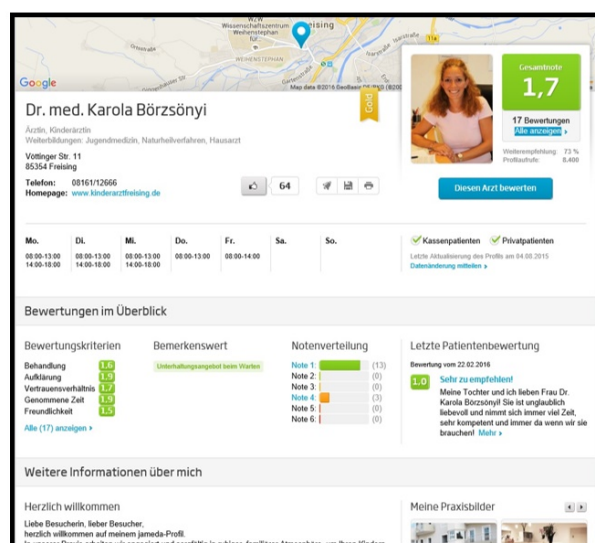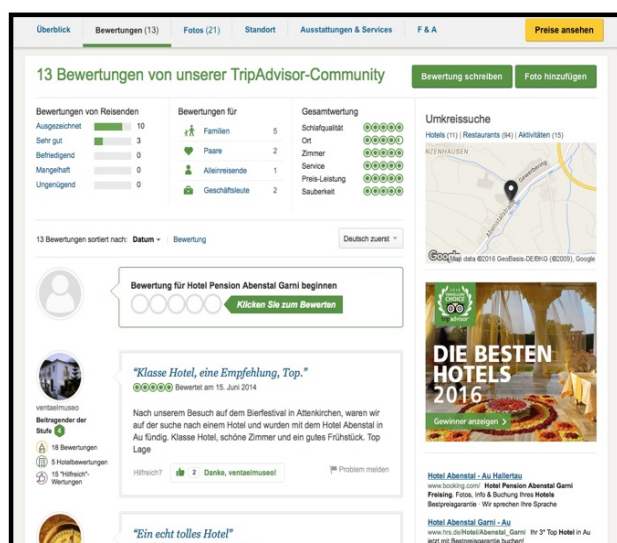
Theoretical Background

Extensive research has been conducted about the users, recipients, search, and decision-making processes on commercial Web-based ratings, reviews, and electronic word of mouth [23]. However, whether or to what extent this research can be translated to user behavior on PRWs (due to a similar setup and design of CRWs and PRWs) still needs to be investigated. How individuals make sense and understand their surroundings has theoretical roots in educational psychology. Resubsumption [24-27] and recategorization theory [27,28] outline how individuals reason when they encounter new phenomena, when they apply the same frame of logic that they learned from a similar earlier experience (monotonic processing) or a different one (nonmonotonic learning). Monotonic processes entail the learning of new information without any changes to the existing knowledge, whereas nonmonotonic learning coincides with changes in cognition, such as in attitudes, beliefs, conceptual or theory change, or deep learning [26,29]. Therefore, when individuals are faced with new situations or information, they by default anchor what they see, read, or perceive monotonically into a scheme that they are already familiar with [25]. In other words, they rely on resubsumption, which is "the process by which an existing theory is considered to be capable of encompassing and explaining additional experiences or phenomena and can occur before any confrontation with anomalies" [27]. In this study, we investigated whether the similar setup and design of CRWs and PRWs led to monotonous processes and subsequently analogous choice making despite the dissimilar attributes of the services that were to be selected (hotel vs physician).

Methods

Study Design and Participant Recruitment

A qualitative approach was employed to provide an in-depth understanding of consumer decision making. We used the consolidated criteria for reporting qualitative research (COREQ) to describe the study methods [30] (Multimedia Appendix 1). On approval of the study protocol by the Ethics Committee of the Università della Svizzera italiana (24.6.2015 CE 2015-5), we performed a total of 22 in-depth interviews with young parents residing in the German-speaking part of Switzerland. Main eligibility criteria included (1) age older than 18 years, (2) fluency in German, and (3) being a parent of a child aged younger than 3 years with no diagnosed physical or mental condition. We used a number of purposive sampling strategies to identify, notify, and invite participants who could cover a wide spectrum of perspectives and experiences: we distributed flyers at venues frequented by the target population (eg, child day care centers, parents' associations) and liaised with midwives visiting young parents at their homes. Snowball techniques were also used, that is, initial participants were asked to pass on information about the study to other people in their networks. Parents interested in participating either gave permission to be contacted or directly got in touch with the research team (via phone) and arranged a convenient date and time for the interview. This multifaceted approach resulted in the recruitment of 22 young parents, out of whom 12 were females. Participant sociodemographic characteristics are presented in Table 1.

Table 1. Participant characteristics.

Variable	Category	n	%
Gender	Male	10	45.5
	Female	12	54.5
Age group	26-30 years	8	36.3
	31-35 years	10	45.5
	36-40 years	4	18.2
Education	University degree	4	18.2
	Applied science university degree	10	45.5
	Apprenticeship	8	36.3
Working status	Full time	9	40.9
	Part time	11	50.0
	Not working	2	9.1
Prior parenting experience[a]	No	15	68.2
	Yes	7	31.8
Self-reported IT proficiency[b]	High	14	63.6
	Medium	4	18.2
	Low	4	18.2

[a]Refers to being a first-time parent or being a parent of an older child at the time of the interview.

[b]As perceived and reported by study participants.

Data Collection

Data for this study were collected from June to August 2015, through one-to-one, face-to-face interviews that lasted about 1 hour. The first author (FR), a social scientist with substantial training in qualitative research, conducted all interviews either in a private room at the University of Lucerne or at participants' homes (based on their individual preferences). At the beginning of the process, the interviewer introduced herself, indicating that she is a PhD student focusing on topics related to how consumers reach their decisions based on Web-based reviews. She then invited informants to complete 2 choice tasks, which involved (1) choosing a hotel based on the commercial Web-based rating website TripAdvisor and (2) selecting a pediatrician based on the PRW, Jameda. To assimilate the search experience for all participants, they were instructed to search for a hotel and a pediatrician in the town of Freising (Germany), whereas the dates of the hotel stay were fixed for all interviewees and the budget restricted to CHF 50-100 per night

(Textbox 1). The costs were automatically displayed in Swiss Francs as the page was accessed from Switzerland. In this way, the number of possible options for hotels and pediatricians was limited to 8 and 7, respectively. To better understand consumers' thought processes, we invited participants to "think aloud" [31], namely to verbalize their thinking while sorting through information and reaching decisions. Using a semistructured interview guide, we subsequently posed open-ended questions to allow them to elaborate more on factors influencing their decision making, level of confidence in their final choice, and perceived differences and similarities in their search for a hotel and a physician. The interview guide was developed on the basis of existing literature and our own interest in the topic; yet, it was not systematically pilot tested. Interviews were held until data saturation was reached. All parents signed an informed consent form before their participation in the study. On completion of the interviews, participants received CHF 20 to compensate for their travel or babysitter expenses.

Textbox 1. Overview of choice tasks

Choice task 1:

Imagine that you are going on a 5-day trip to Freising near Munich, Germany, with your family. You are leaving on November 2, 2015, and you are coming back on November 7, 2015. Please go on tripadvisor.de (the first tab in the browser window that is already open on the laptop) to choose a hotel that fits your needs. Please keep in mind that your budget per night is between 50-100 Francs, so please do not modify this criterion. You have about 15 minutes to complete this task.

Choice task 2:

Imagine that you or your partner had to change jobs and you had to move to Freising near Munich, Germany. Once you move there, your child has high fever and you urgently need to see a pediatrician. You have not met anyone yet in Freising, so you decide to consult the Internet. Please go to jameda.de (the second tab on your laptop browser window) to choose a pediatrician for your child. You have about 15 minutes to complete this task.

Analysis

With participant permission, all interviews were recorded, using a digital voice recorder, and fully transcribed, along with field notes taken during and/or after the process. Data were subsequently anonymized and coded using the qualitative data analysis software Atlas-Ti. We did not use any pre-existing coding frame; rather, we opted for identifying themes linked to the data themselves. Using the approach by Braun and Clarke to thematic analysis [32], data analysis was a multistage, recursive process, which involved moving back and forth throughout the following phases: (1) familiarization with the data; (2) generating initial codes; (3) searching for themes; (4) reviewing themes; and (5) defining and naming themes [32]. Analysis was performed on the original Swiss German transcripts by FR, who is a native Swiss German speaker, with a subsample translated in English and being analyzed independently by EG to verify emergent themes.

Results

Participants spent on average 9:57 minutes (standard deviation=9:22, minimum=3:46, maximum=22:25) searching for a hotel and 6:17 minutes (standard deviation=4:47, minimum=00:38, maximum=19:25) for a pediatrician. In fact, 19 of the 22 participants took longer to choose a hotel than a physician (see Figure 2). When choosing an accommodation, individuals stated to have specific criteria in mind that were based on their past experience with hotels and review websites. Therefore, parents strategically went searching for information about the location and accessibility of the hotel, the hygiene standards, the availability of children-friendly rooms and appliances, and the sleep quality. This took them considerable search time because finding such information was not always straightforward, especially if participants were not familiar with the TripAdvisor website.

By comparison, when choosing a pediatrician, individuals did not know exactly what to look for because all (except one) reported never having visited or heard of PRWs. Yet, due to the similar layout of the websites (Figure 1), almost all participants initially looked for the same type of information as they did when searching for a hotel. For example, individuals who had focused on pictures in their hotel selection on the CRW also checked out physicians' photos first on the PRW. Likewise, participants who had been keen on reviews for hotels looked right away at comments when selecting a physician. One gender difference presented itself in the searches: although the large majority of men applied filters in their searches for both hotels and pediatricians (which allowed them to specify their criteria and cut down their choices), only a small minority of women did so. A male participant described the two search tasks:

> It was very similar. Selecting what I am looking for, applying the filters accordingly, and then I started from the top. After that, I looked at the reviews and then some pictures. That's important because these are the first things you look at. (…) It was quite similar. [BB, male, 34 years]

Although the approach for choosing a hotel and a physician was highly similar at first sight, the interviews following the search tasks displayed major differences between the use of CRWs and PRWs. Four main themes that emerged from the analysis of our interview data can explain the differences between the two search tasks: (1) trial and error, (2) trust, (3) competence assessment, and (4) affect and likeability. The 4 themes were often brought up in an interwoven manner. Frequently one theme led to another in explaining the decision process and the reasons why physician choice was easier and carried more confidence than the hotel selection. Partial overlap of the themes was therefore inevitable.

Figure 2. Total search time by task and participant.
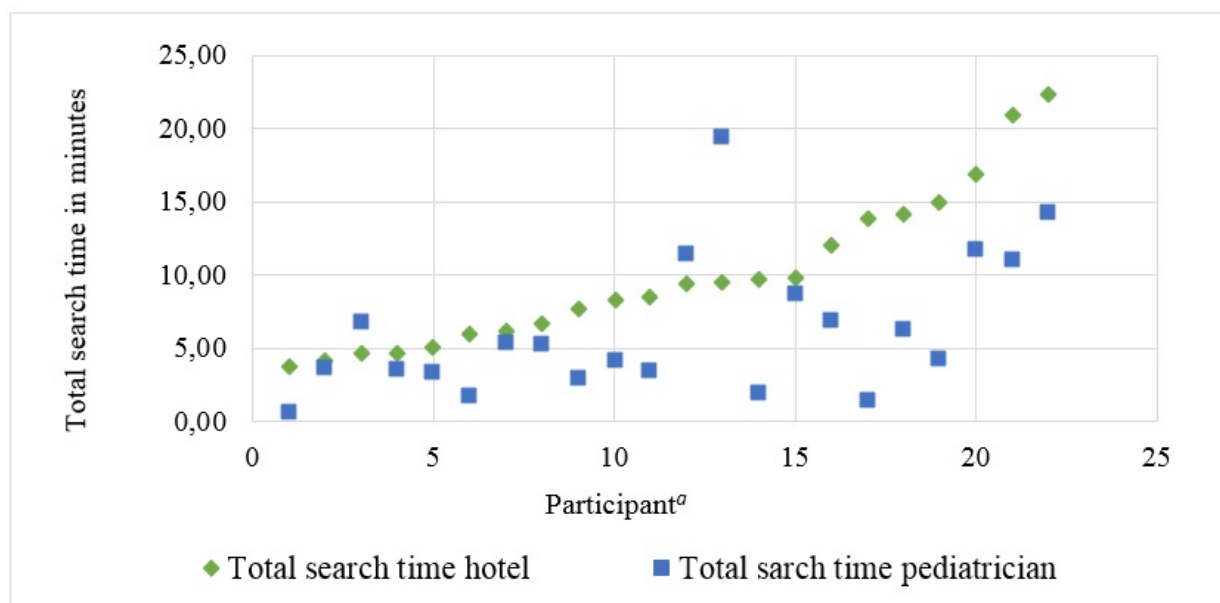
ᵃParticipants are listed in ascending order, according to their hotel search time.

Theme 1: Trial and Error

In contrast to the choice of an appropriate accommodation, where individuals searched for detailed criteria that were important for them, parents unanimously applied a "trial and error" approach when choosing a pediatrician. Participants stated that, regardless of the kind or quality of information provided on PRWs, they would only know after the first visit whether they felt a connection with a physician or not. Hence, they would try a first consultation and subsequently decide to stay with that pediatrician or to switch to another doctor:

> *I basically trust all physicians and if I had visited a physician and somehow felt like "no, this wasn't quite right," then I would maybe search again.* [AB, female, 30 years]

> *I would go and give him a try and if it doesn't work or doesn't fit, I would change.* [SKG, female, 34 years]

Participants justified this "trial and error" approach by stating that the costs and stakes for choosing a pediatrician and a hotel were unequal. Although the choice of a physician was described as more important, the felt pressure to choose an appropriate accommodation seemed higher. Participants explained that while they had to spend their own money on accommodation, pediatrician visits were typically covered by the health insurance. This meant that the financial costs at stake were higher for the hotel selection, thus the choice of hotels had to be made more thoroughly. In addition, participants voiced profound trust in the medical profession, and hence, they perceived any of the listed doctors to be competent to treat their child (see theme 2). Therefore, individuals could afford to check out a pediatrician and then consult someone else if they were not satisfied. Furthermore, individuals were to invest scarce free time in an accommodation of their choice, whereas pediatricians visit would be of short duration and, therefore,

would require a minor time commitment. Therefore, participants felt less pressured to right away make an optimal choice of pediatrician, which subsequently shortened their physician search effort and time compared with a hotel.

> *The hotel is more about something like holidays, something nice and the physician is a necessity. It isn't somewhere/something where you go to relax. That's why there is a difference. I think with a physician, your insurance actually pays most of it. And with the hotel you have to pay but you can also freely choose and that's why you want a good value for your money.* [AB, female, 30 years]

The "trial and error" approach was also explained by the usage behavior of the different websites. PRWs could serve as a way to get an initial overview of the available doctors in an area, to make a first contact with a physician. The hotel rating website on the other hand was expected to provide a set of specific information that allows individuals to reach a final choice or to book an accommodation on the spot. Individuals compared the PRWs with an improved version of a phone book, in which they would otherwise look up the physicians' contact information. Even though Web-based PRWs were considered as a better information source than a phone book, the additional information on the website was still perceived as insufficient to make a definite pediatrician choice. On the other hand, with the hotel choice, they felt confident to make a definite choice based on a rating website like TripAdvisor.

> *Yes, it is quite difficult but the website provides you at least with more information than the phone book, which would be the alternative. At least that's how it (pediatrician choice based on the phone book) was done when there were no alternatives yet. In that sense, it is still a gut feeling decision in the end: I don't know what will await me but I would just have to try once. For example with a physician, if I wasn't*

satisfied, I could go to someone else the next time around. [SI, female, 33 years]

The different website consumption was also based on the awareness that the choice of hotel would be definite and nonreversible on arrival, whereas a first visit at a pediatrician was not binding. Therefore, a trial and error approach seemed affordable with a physician but not with a hotel:

> *Such a physician rating website would probably be useful to make a first contact but after that it is obviously very much about the feeling you get, the appearance and impression once you get there. With a hotel, you book and then you say afterwards "okay, that was great" and you may go again some other time. There it is about the best offer at that moment. It isn't really that relevant.* [LS, female, 27 years]

Theme 2: Trust

Although most participants searched for technical skill information about the doctors, a high-level respect toward physicians, and, more broadly, trust in the health care system was evident in participants' accounts. They tried to find descriptive information about the qualifications and certifications of doctors because the information seemed more legitimate to them. Hence, they did not question the information provided by the website provider or the physician himself. On the other hand, reviews and ratings by former patients were often perceived as strange or even unacceptable; participants could hardly believe nor approve that the skills of a physician would be evaluated by lay people. This suspicion toward physician ratings led some individuals to not even look at available reviews. This cut down their search time because subsequently they would choose a doctor based on sociodemographic indicators, location or decision shortcuts, such as to take the first physician on the list or to select the doctor closest to their home. One mother described:

> *This is just a bit strange to me, this whole thing about ratings! Especially rating a physician –to me, this is just suspicious! Yeah, I just still feel like: it is a physician! (…) Well, I don't know, I would probably just take the one that is closest to my house.* [AB, female, 30 years]

In contrast to the choice of pediatrician, participants' confidence in hotels was lower. Accommodations of the same star ratings (certificates by an external expert source) for example, were thought to differ tremendously in terms of quality. This, in turn, increased the time they spent searching for a hotel, while it decreased their choice confidence. Participants expressed their trust in physicians by referencing the well-known and rigorous educational requirements that all certified doctors in Switzerland and Germany have to fulfill. Parents concluded that all physicians needed to pass a critical threshold of competence because, otherwise, they would not have been certified as medical doctors. All listed doctors would therefore be capable and able to help their child. Hence, the difference in technical competence among pediatricians was perceived to be minor or nonsignificant, whereas the quality of hotels despite equal star ratings could differ tremendously.

I always think: they are physicians! And I actually always have a positive attitude anyways. (…) that's why I am just glad if someone is there who can help if there is a need. [SA, female, 30 years]

I just think: "a physician is a physician." Somehow I am just... They know what they are doing! [AB, female, 30 years]

Theme 3: Competence Assessment

Although participants displayed confidence and experience in evaluating the quality of a hotel, they perceived themselves less competent in assessing the quality of a physician. Due to this perceived inability to assess the quality of a physician and the subsequent helplessness that they felt, the vast majority of parents opted to trust physicians by default. One mother explained:

> *I don't understand any of it (the diagnosis and treatment prescription), so I trust in that, what he tells me and then I just take that (the medication) and I do what he tells me to do.* [EA, female, 34 years]

> *Obviously there are differences (among physicians), it always depends upon what your problem is. But in the end... Yeah well, you also don't know which one (pediatrician) is better than the other ones. You never know!* [EA, female, 34 years]

The perceived incompetence of health consumers to correctly evaluate the skills and abilities of a physician was not only restricted to one's own experience of a consultation with a particular physician but individuals also attributed it to former patients/reviewers who appraised the competence of a pediatrician on PRWs. As one mother states, PRW reviews should be taken with a grain of salt because according to her, a regular patient, a parent of a patient, or she herself would not be capable of assessing physician competence:

> *"A very good physician" – who am I to evaluate that?!I cannot assess the diagnosis, I am not a doctor! Under most circumstances I cannot tell whether it (the diagnosis and treatment) was good or not. Just based on gut feeling, this one (pediatrician) was likeable to me. But that doesn't help me here... she (this physician) can be super nice and thoroughly answer to all questions, but in the end, she could be telling the biggest rubbish and prescribe a medicine that doesn't fit the actual disease symptoms. I just cannot be sure.* [SI, female, 33 years]

Parents found that there are more criteria that they could consider in their choice of hotels than those in the choice of physicians. This lack of knowledge about what makes a good physician and the resulting trust when realizing that one is not able to differentiate was not mirrored in the choice of hotels. Participants voiced confidence in their hotel assessment expertise. This is the reason why individuals' search for an accommodation was more detailed, more structured, and focused:

> *A hotel rating website probably provides you with more criteria to choose from, which makes the decision more difficult. And here (referring to the*

PRW), I mean you have seven pediatricians. They are all somewhat quite similar, grades between 1 and 2.3 and I mean, there isn't any one that is somehow very, very bad. [FG, male, 30 years]

The general skepticism toward reviewers (as displayed in the citations previously) was dominant for reviews on PRWs. On the other hand, reviews by former hotel guests were given more weight and considered in more depth, whereas the reviews of physicians were perceived to be more situation dependent and therefore most likely less representative than for hotels. By stating this, individuals implied that situation-dependent reviews are not an indicator of the actual and true qualities of a physician. Reviews may be biased due to the reviewer's personal involvement, and therefore, a trial and error approach would be necessary to create a personal opinion. One mother described this difference:

I find it difficult to give a grade to a doctor. For a hotel room it is much easier than for a physician because how you experience a physician is extremely dependent on the time, the day, and many other things. Yes. I think if you get unlucky, yeah... And many positive ones don't even write reviews. But if you are angry, then you go onto the website and you say: I want to hurt this person with a bad review, yeah. [SA, female, 30 years]

Theme 4: Affect and Likeability

Informants often let the emotions that appraisals or pictures of the physician or hotel evoked guide their choice making. Both positive affect to get a good sensation about a choice option and negative affect that led individuals to turn down or avoid an option were present. However, participants stated that the choice of a pediatrician was much more emotion guided than the selection of a hotel. Interpersonal connection, gut feeling, and likeability played a large role in the choice of a doctor, whereas the commercial focus on price and offered facilities was crucial for the choice of hotels. As one father, who was in favor of reviews, described:

With a hotel it was much more important how much it costs and this component isn't present for a physician because it will most likely be paid by the insurance. But other than that, they are very similar, especially at the beginning. Afterwards, I think the hotel is more about commercial aspects, while the physician is certainly more about feelings and that's why the reviews from other people are much more important. [LS, male, 38 years]

The likeability and connection with a physician was also described to be very subjective and highly dependent on the individual. As personal affect and likeability play an important role in the choice of a physician, participants would often admit to apply a trial and error approach based on their own experience because reviews or word of mouth did not seem representative:

One may say "he is an excellent doctor, you have to go to him!" and someone else says "oh no, I don't agree at all". This is the same for a midwife and for

a dentist. You have to see it for yourself, if it fits for you and if you like it. [EA, female, 34 years]

The first positive affect-evoking aspect was the perception of likeability of the physician. A variety of indicators were used, in particular, reviewers' evaluations of personality aspects and sociodemographic information. The former included criteria, such as a physician's sense of humor, the likeability, and the perceived child friendliness:

I find likeability, charisma, how he treats us, how he treats our child, how he talks to us, very important. [FG, male, 31 years]

Sociodemographic factors typically consisted of the pediatrician's age, the gender, or the perception that the pediatrician may share common interests with the parent. In addition, the cultural background was inferred from the pediatrician's name, with doctors with German-sounding last names being preferred. Participants stated that knowing about a physician as a person gave them a positive gut feeling and that shared characteristics made them feel more confident about the first encounter:

I think physicians don't have to hide behind these reviews and it would have just been nice if there would have been some more information about them or a picture of the practice maybe. That way you can already get a good feeling about where you are going for the visit and some trust is already built through the Internet I would say. [AW, female, 29 years]

Furthermore, pictures were often used to make an inference about a person's likability:

The picture is the most important indicator for me: is this person likeable or not? I am someone who trusts in the first impression. Obviously there is a possibility to convince me of the opposite but I can't deny that if someone is likeable at first sight, then I can just go much more relaxed and comfortable to that visit than if someone isn't. [RS, male, 32 years]

In addition to aspects concerning the character of a physician, individuals were also interested in the beliefs that a pediatrician held. In particular, the attitude toward vaccination and alternative medicine posed key decision factors. Especially parents who valued alternative medicine and had a critical attitude toward vaccination tried to find information about pediatricians' views on those topics. Once they found a pediatrician who shared their beliefs, it had a large impact on their decision as it increased confidence and reduced search time:

(...) if you don' want to get your child vaccinated, then you have to really think hard which physician you want to consult, so that you can enforce that once you get there. [PS, male, 33 years]

Although individuals tried to find a physician who evoked positive emotions, certain factors led them to turn away immediately. For hotel choice, negative affect was caused by pictures, a lack of child-friendly appliances or furniture, or information about the lack of hygiene. For physician choice, on the other hand, insufficient office organization or planning

capacity and a lacking of willingness to communicate and inform parents about the treatment caused the arousal of negative emotion. When looking at reviews, physicians were judged in a milder way than hotels because the reviews were perceived to be more subjective and therefore of relative importance. Hence, individuals were less concerned about negative physician reviews, whereas negative comments about hotels were given much more weight:

> A very nice doctor" and the next person states that he was in a bad mood. For me, these are not hard facts like you can find them for a hotel. With a hotel it certainly is subjective too but still, it is not quite the same for me. [SI, female, 33 years]

> That's with every physician - they are also people and everyone has a bad day sometimes and most of the time it's exactly then that people spread the word and harp on about it. [SA, female, 30 years]

Discussion

Our results offer insights into consumer decision making based on Web-based reviews. At first sight, participants in this study seemed to apply the same search strategy and decision mechanisms when choosing an accommodation and when selecting a pediatrician; on both websites, each individual focused on the same type of information to reach a decision, such as pictures of the rooms or the practice, reviews, or descriptive information about facilities and services or treatments offered. The highly similar layout of the two websites led them to start out in the same manner in the hotel search as in the physician selection. Yet, as the search proceeded, individuals became more skeptical and changed their approach through the discovery of differences between the choice of an accommodation and the choice of a physician. In the interviews after the choice tasks, participants' conceptualization of the two websites appeared entirely different. The themes that emerged from our interviews describe this sense-making process or, consistent with resubsumption [24,25] and recategorization theory [27,29], the change from monotonic processing (both websites are the same and should be approached equally) to nonmonotonic change (the differences between the 2 choices outweigh the similarities and therefore the decision-making process for the 2 choice changes).

Overall, our results suggest that, although the choice of a pediatrician was perceived as more important than the choice of a hotel, participants found choosing a physician much easier than selecting a suitable accommodation. Moreover, although participating young parents seemed to believe that their choice of hotel could have been ameliorated, they were more satisfied with their choice of pediatrician, despite the shorter search time. On the basis of our interview data, these discrepancies can be explained by the following factors: (1) the application of a "trial and error" approach in the selection of an appropriate medical doctor; (2) their high trust in medical expertise and the Swiss health care system; (3) participants' perceived inability to properly evaluate the skills and abilities of physicians; and (4) the role of "gut feeling" and likeability (vs the use of explicit criteria) in choosing a physician.

A starting point to understand our findings lies in the economics literature, which has developed a typology of service goods, categorized according to the information asymmetry between provider/seller and customer/consumer [33,34]. The first type consists of the so-called "search goods." They can be compared before purchase in terms of price and quality based on openly available information, which in turn leads to high levels of choice confidence [35]. Research has found that for search goods, the risk perception is lowest, whereas the search time (due to the strategic comparison of the goods) is highest unless the cost of searching would outweigh the price of the good [35]. For the second type, namely "experience goods," information inequality is larger. Consumers are unable to assess the quality of this kind of product or service until tried, consumed, or experienced [34]. Therefore, search time is shorter until a decision is reached because a trial of the product is essential to evaluate the quality of experience goods [35]. For the third type, that is, "credence goods," the information inequality between provider and consumer is the highest. Because individuals cannot assess the quality of the product even after consumption, the search time is lower and individuals rely more on interpersonal recommendations than public noncustomized information before purchase or selection [35].

Individuals' description of the search process for an accommodation and a pediatrician was reflected in the categorization of the goods and services categorization above. Although the selection of hotels mainly carried attributes of an experience good, the choice of a physician could predominantly be classified as credence good. As participants in this study described, they were highly motivated to compare possible options for accommodation, as they would have to invest both their own money and their holiday time. In addition, they felt confident that, based on the information provided, they could choose a good hotel. The vast majority of participants also appeared to trust in the information provided on TripAdvisor and meticulously compared the available hotel choices. Yet, informants stated that they could only know if they would be satisfied with their choice once they experienced a stay at the chosen accommodation.

For the choice of the physician, on the other hand, individuals decided faster and described that they applied a trial and error approach. Participants often perceived that using a PRW would only marginally aid their choice and they preferred making an appointment and meeting the physician in person. Participants were sure that after a first visit, they would be able to understand if they had a good interpersonal connection with the doctor. However, they were aware that likability was not an indication of the quality of the physician or the correctness of the diagnosis and treatment they received. Therefore, individuals' description of the Web-based physician choice resembled the attributes of a credence service; evaluation of the quality of treatment could almost never be assessed even after an appointment, neither by themselves nor by other laypeople who posted comments or reviews on PRWs. Hence, many participants were skeptical about comparing physicians, as previous research had suggested [36].

Despite the congruent findings between this study and existing literature on search, experience, and credence goods, there are

some theoretical expansions that the results of this study suggest. First, the theory on good classification according to Nelson [33] poses that the trust into one's own choice should be highest for search goods and diminish to be the lowest for credence goods. Nevertheless, our study suggests that in the context of health care, this may not hold true. Due to the high trust in physicians that individuals of this study held, the perceived riskiness of a credence good and the confidence in one's choice is dependent on the context. Not only the attributes of the good or service but also an individual's past experience, and the level of trust in the system in which that good is provided, should be taken into consideration. Trust may be a moderator or mediator in the search and choice making of credence good.

Limitations

This study has certain limitations that need to be addressed. First, our results should be interpreted with caution, given that participants' accounts were triggered by 2 hypothetical scenarios and not real-life occurrences (eg, an actual move of the family to a foreign place or a real illness of the child). Yet, we opted for using real websites in an effort to create a more realistic sense and enable informants to reflect on their prior experiences. Nevertheless, this also meant that we were not able to control for any modifications caused by the dynamic and volatile nature of these websites. In fact, the information that individuals viewed in the 2 choice tasks slightly differed from the start of the data collection (June 2015) toward the end (August 2015) because of newly posted comments and reviews. In order, however, to maximize the information that all participants would see and could base their choice on, price category for the hotel and visiting dates were fixed.

Second, the study sample came from the German-speaking part of Switzerland, in which education levels, health literacy, and numeracy have been found to be very high [37]. Previous research has shown that active search of physicians and correct evaluation of quality of care information to select a physician depended on these 2 characteristics [38-40]. Therefore, although our findings could be transferable to other contexts characterized by high health literacy and numeracy levels, they might be less relevant for contexts where these attributes are less pronounced.

Third, given the context-bound nature of qualitative research, our results should be interpreted in relation to the specific legal and factual circumstances of Switzerland. Specifically, Switzerland has not yet established a common electronic health record, which is state of the art for many other European countries. Apart from this, the legal situation for Swiss PRWs differs from its European neighbors. Calling on data protection laws, the Swiss physicians' association (FMH) requested a ban of negative reviews for physicians practicing in Switzerland [5]. This has affected the dissemination and use of PRWs in the country, which can explain the low awareness of PRWs among study participants.

Conclusions

Individuals' page navigation on PRWs and CRWs appeared similar. However, their perception of the 2 choice tasks was different. Participants displayed greater confidence in their physician selection, but took less time to make that choice. Individuals were incapable to assess whether a physician was good or not, even if they reviewed available information on the review website. This led them to a "trial and error" approach; only after a first encounter, they would decide whether the doctor was a match in terms of likeability, yet they would still remain unsure about his or her technical qualities. On the other hand, for the hotel choice, participants were confident that they could obtain all necessary information on the Internet to choose a good hotel for themselves. The fact that they invested their own money and precious free time made them feel more responsible to make a good hotel choice. Therefore, the information provided on CRWs seemed to fit customers' needs, whereas the similarly designed PRWs did not satisfy study participants due to the attributes of the 2 services in the choice tasks.

Implications for Future Research

In this study, participants took less time to search for physicians than hotels, voiced a profound trust in the health care system, and applied a "trial and error" approach to choose a doctor. Previous research has found that particularly in developed health care systems, such as Switzerland or Germany, patients refrain from actively searching and comparing physicians [41]. Therefore, a replication of this study in a country where trust in the health care system is lower may bring additional valuable insights. PRWs and CRWs look very similar: both offer pictures, background information (location, contact information, and descriptions of the facilities and services offered), as well as numeric written reviews by former guests or patients. However, as this study showed, the 2 service goods cannot be treated equally because of their unequal attributes. This has implications for the information needs consumers or patients have and, subsequently, for the design and creation of PRWs. Although individuals seem to want mass media information sources for search goods, a combination of mass media and interpersonal information for experience goods, they rely heavily on their social network to obtain information for credence good choices [35]. Currently, PRWs are set up in the same manner as an experience good by combining mass media information (general information about the location, accessibility, qualifications, and so forth) of the physician with impersonal recommendations (such as anonymous reviews by former patients). Interpersonal information that is first of all trustworthy, and second, adjusted to individuals' needs [35,42] is therefore required for credence goods. Hence, this study provides first indications to change the way PRWs are currently designed. Translating these requirements into a website would be a task for today's Web designers and researchers to take on. In addition, comparisons between the choice making process for a physician and other experience and credence goods/services (such as for a plumber, hairdresser, or car mechanic) could provide further insights into users' product or service-based information needs. The results of this study could serve as the basis for future research that focuses on individuals' information needs when selecting a doctor on the Internet. Experimental tests could subsequently improve the design and, therefore, the use and navigation of Web-based PRWs.

Acknowledgments

Authors' Contributions

The three authors jointly conceived and designed the study. FR recruited the sample, collected the data, performed data analysis, and wrote the initial draft. EG contributed to data analysis and critically revised the draft. PJS provided expert comments on the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

COREQ checklist.

[PDF File (Adobe PDF File), 210KB - jmir_v18i6e129_app1.pdf]

References

1. Hennig-Thurau T, Gwinner KP, Walsh G, Gremler DD. Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? Journal of Interactive Marketing 2004 Jan;18(1):38-52. [doi: 10.1002/dir.10073]
2. Emmert M, Sander U, Esslinger AS, Maryschok M, Schöffski O. Public reporting in Germany: the content of physician rating websites. Methods Inf Med 2012;51(2):112-120. [doi: 10.3414/ME11-01-0045] [Medline: 22101427]
3. Hanauer DA, Zheng K, Singer DC, Gebremariam A, Davis MM. Public awareness, perception, and use of online physician rating sites. JAMA 2014 Feb 19;311(7):734-735. [doi: 10.1001/jama.2013.283194] [Medline: 24549555]
4. Emmert M, Meier F, Pisch F, Sander U. Physician choice making and characteristics associated with using physician-rating websites: cross-sectional study. J Med Internet Res 2013;15(8):e187 [FREE Full text] [doi: 10.2196/jmir.2702] [Medline: 23985220]
5. Atasoy S. SRF Schweizer Radio und Fernsehen online. 2014 Nov 24. Patienten benoten Schweizer Ärzte URL: http://www. srf.ch/news/schweiz/patienten-benoten-schweizer-aerzte [accessed 2016-03-17] [WebCite Cache ID 6g4cbacYc]
6. Emmert M, Meier F. An analysis of online evaluations on a physician rating website: evidence from a German public reporting instrument. J Med Internet Res 2013 Aug;15(8):e157 [FREE Full text] [doi: 10.2196/jmir.2655] [Medline: 23919987]
7. 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]
8. Hanauer DA, Zheng K, Singer DC, Gebremariam A, Davis MM. Parental awareness and use of online physician rating sites. Pediatrics 2014 Oct;134(4):e966-e975 [FREE Full text] [doi: 10.1542/peds.2014-0681] [Medline: 25246629]
9. McBride DL. Parental use of online physician rating sites. J Pediatr Nurs 2015 Jan;30(1):268-269. [doi: 10.1016/j.pedn.2014.10.003] [Medline: 25450443]
10. Strech D. Ethical principles for physician rating sites. J Med Internet Res 2011 Dec;13(4):e113 [FREE Full text] [doi: 10.2196/jmir.1899] [Medline: 22146737]
11. Greaves F, Pape UJ, King D, Darzi A, Majeed A, Wachter RM, et al. Associations between Internet-based patient ratings and conventional surveys of patient experience in the English NHS: an observational study. BMJ Qual Saf 2012 Jul;21(7):600-605. [doi: 10.1136/bmjqs-2012-000906] [Medline: 22523318]
12. McCartney M. Will doctor rating sites improve the quality of care? No. BMJ 2009;338:b1033. [Medline: 19293224]
13. Lagu T, Hannon NS, Rothberg MB, Lindenauer PK. Patients' evaluations of health care providers in the era of social networking: an analysis of physician-rating websites. J Gen Intern Med 2010 Sep;25(9):942-946 [FREE Full text] [doi: 10.1007/s11606-010-1383-0] [Medline: 20464523]
14. Ma L, Kaye AD, Bean M, Vo N, Ruan X. A five-star doctor? Online rating of physicians by patients in an internet driven world. Pain Physician 2015;18(1):E15-E18 [FREE Full text] [Medline: 25675065]
15. Grabner-Kräuter S, Waiguny Martin K J. Insights into the impact of online physician reviews on patients' decision making: randomized experiment. J Med Internet Res 2015 Apr;17(4):e93 [FREE Full text] [doi: 10.2196/jmir.3991] [Medline: 25862516]

XSL•FO

RenderX

16. Lewis P, Kobayashi E, Gupta S. An online review of plastic surgeons in southern California. Ann Plast Surg 2015 May;74 Suppl 1:S66-S70. [doi: 10.1097/SAP.0000000000000517] [Medline: 25875912]

17. Greaves F, Pape UJ, Lee H, Smith DM, Darzi A, Majeed A, et al. Patients' ratings of family physician practices on the internet: usage and associations with conventional measures of quality in the English National Health Service. J Med Internet Res 2012;14(5):e146 [FREE Full text] [doi: 10.2196/jmir.2280] [Medline: 23076301]

18. Gilbert K, Hawkins CM, Hughes DR, Patel K, Gogia N, Sekhar A, et al. Physician rating websites: do radiologists have an online presence? J Am Coll Radiol 2015 Aug;12(8):867-871. [doi: 10.1016/j.jacr.2015.03.039] [Medline: 26250980]

19. Samora JB, Lifchez SD, Blazar PE. Physician-Rating Web Sites: Ethical Implications. J Hand Surg Am 2016 Jan;41(1):104-110.e1. [doi: 10.1016/j.jhsa.2015.05.034] [Medline: 26304734]

20. Emmert M, Meier F, Heider A, Dürr C, Sander U. What do patients say about their physicians? an analysis of 3000 narrative comments posted on a German physician rating website. Health Policy 2014 Oct;118(1):66-73. [doi: 10.1016/j.healthpol.2014.04.015] [Medline: 24836021]

21. TripAdvisor. Hotel Pension Abenstal Garni URL: https://www.tripadvisor.ch/Hotel_Review-g198549-d1891130-Reviews-Hotel_Pension_Abenstal_Garni-Freising_Upper_Bavaria_Bavaria.html [WebCite Cache ID 6fhhKx17D]

22. Jameda: Deutschlands grösste Artzempfehlung. Physician profile of dr. Karola Boerzsoenyi URL: http://www.jameda.de/freising/aerzte/kinderaerzte/dr-karola-boerzsoenyi/uebersicht/81074126_1/ [WebCite Cache ID 6fhg9uTWV]

23. King RA, Racherla P, Bush VD. What we know and don't know about online word-of-mouth: a review and synthesis of the literature. Journal of Interactive Marketing 2014 Aug;28(3):167-183. [doi: 10.1016/j.intmar.2014.02.001]

24. OHLSSON S. Resubsumption: a possible mechanism for conceptual change and belief revision. Educational Psychologist 2009 Jan 23;44(1):20-40. [doi: 10.1080/00461520802616267]

25. Ohlsson S. Meaning change, multiple routes, and the role of differentiation in conceptual change: alternatives to resubsumption? Educational Psychologist 2009 Jan 23;44(1):64-71. [doi: 10.1080/00461520802616309]

26. Ohlsson S. Deep learning: how the mind overrides experience. New York, USA: Cambridge University Press; 2011.

27. Chi MTH, Brem SK. Contrasting Ohlsson's resubsumption theory with Chi's categorical shift theory. Educational Psychologist 2009 Jan 23;44(1):58-63. [doi: 10.1080/00461520802616283]

28. Chi MTH. Three types of conceptual change: belief revision, mental model transformation, categorical shift. In: Vosniadou S, editor. International Handbook of Research on Conceptual Change. New York, USA: Routledge; 2008:61-82.

29. Ramsburg J, Ohlsson S. Category change in the absence of cognitive conflict. J Educ Psychol 2016;108(1):98-113.

30. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]

31. Van SM, Barnard Y, Sandberg J. The think aloud method: A practical guide to modelling cognitive processes. London: Academic Press; 1994.

32. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101.

33. Nelson P. Information and consumer behavior. J Polit Econ 1970;78(2):311-329.

34. Hsieh Y, Chiu H, Chiang M. Maintaining a committed online customer: A study across search-experience-credence products. Journal of Retailing 2005 Jan;81(1):75-82. [doi: 10.1016/j.jretai.2005.01.006]

35. Mitra K, Reiss MC, Capella LM. An examination of perceived risk, information search and behavioral intentions in search, experience and credence services. Journal of Services Marketing 1999 Jun;13(3):208-228. [doi: 10.1108/08876049910273763]

36. Lupton D, Donaldson C, Lloyd P. Caveat emptor or blissful ignorance? Patients and the consumerist ethos. Soc Sci Med 1991;33(5):559-568. [Medline: 1962227]

37. Connor M, Mantwill S, Schulz PJ. Functional health literacy in Switzerland - validation of a German, Italian, and French health literacy test. Patient Educ Couns 2013 Jan;90(1):12-17. [doi: 10.1016/j.pec.2012.08.018] [Medline: 23089240]

38. Faber M, Bosch M, Wollersheim H, Leatherman S, Grol R. Public reporting in health care: how do consumers use quality-of-care information? A systematic review. Med Care 2009 Jan;47(1):1-8. [doi: 10.1097/MLR.0b013e3181808bb5] [Medline: 19106724]

39. Peters E, Dieckmann N, Dixon A, Hibbard JH, Mertz CK. Less is more in presenting quality information to consumers. Med Care Res Rev 2007 Apr;64(2):169-190. [doi: 10.1177/10775587070640020301] [Medline: 17406019]

40. Victoor A, Delnoij DMJ, Friele RD, Rademakers JJDJM. Determinants of patient choice of healthcare providers: a scoping review. BMC Health Serv Res 2012;12:272 [FREE Full text] [doi: 10.1186/1472-6963-12-272] [Medline: 22913549]

41. Fotaki M, Roland M, Boyd A, McDonald R, Scheaff R, Smith L. What benefits will choice bring to patients? Literature review and assessment of implications. J Health Serv Res Policy 2008 Jul;13(3):178-184. [doi: 10.1258/jhsrp.2008.007163] [Medline: 18573768]

42. Arndt J. Word of mouth advertising and information communication. In: Cox D, editor. Risk Taking and Information Handling in Consumer Behavior. Boston, MA: Division of Research, Graduate School of Business Administration, Harvard University; 1967:188-239.

43.   Kanouse DE, Schlesinger M, Shaller D, Martino SC, Rybowski L. How Patient Comments Affect Consumers' Use of
        Physician Performance Measures. Med Care 2016 Jan;54(1):24-31. [doi: 10.1097/MLR.0000000000000443] [Medline:
        26551765]
44.   Schlesinger M, Kanouse DE, Martino SC, Shaller D, Rybowski L. Complexity, public reporting, and choice of doctors: a
        look inside the blackest box of consumer behavior. Med Care Res Rev 2014 Oct;71(5 Suppl):38S-64S. [doi:
        10.1177/1077558713496321] [Medline: 23999489]

Abbreviations

CRW: commercial rating website
PRW: physician rating website

XSL•FO
RenderX

Original Paper

The Impact of an Online Crowdsourcing Diagnostic Tool on Health Care Utilization: A Case Study Using a Novel Approach to Retrospective Claims Analysis

Jessie L Juusola[1], PhD; Thomas R Quisel[1], BS; Luca Foschini[1], PhD; Joseph A Ladapo[2], MD, PhD

[1]Evidation Health, Inc, San Mateo, CA, United States

[2]New York University School of Medicine, Departments of Medicine and Population Health, New York, NY, United States

Corresponding Author:
Jessie L Juusola, PhD
Evidation Health, Inc
15 N. Ellsworth Ave
San Mateo, CA, 94401
United States
Phone: 1 650 279 8855
Fax: 1 650 279 8855
Email: jjuusola@evidation.com

Abstract

Background: Patients with difficult medical cases often remain undiagnosed despite visiting multiple physicians. A new online platform, CrowdMed, uses crowdsourcing to quickly and efficiently reach an accurate diagnosis for these patients.

Objective: This study sought to evaluate whether CrowdMed decreased health care utilization for patients who have used the service.

Methods: Novel, electronic methods of patient recruitment and data collection were utilized. Patients who completed cases on CrowdMed's platform between July 2014 and April 2015 were recruited for the study via email and screened via an online survey. After providing eConsent, participants provided identifying information used to access their medical claims data, which was retrieved through a third-party web application program interface (API). Utilization metrics including frequency of provider visits and medical charges were compared pre- and post-case resolution to assess the impact of resolving a case on CrowdMed.

Results: Of 45 CrowdMed users who completed the study survey, comprehensive claims data was available via API for 13 participants, who made up the final enrolled sample. There were a total of 221 health care provider visits collected for the study participants, with service dates ranging from September 2013 to July 2015. Frequency of provider visits was significantly lower after resolution of a case on CrowdMed (mean of 1.07 visits per month pre-resolution vs. 0.65 visits per month post-resolution, $P=.01$). Medical charges were also significantly lower after case resolution (mean of US $719.70 per month pre-resolution vs. US $516.79 per month post-resolution, $P=.03$). There was no significant relationship between study results and disease onset date, and there was no evidence of regression to the mean influencing results.

Conclusions: This study employed technology-enabled methods to demonstrate that patients who used CrowdMed had lower health care utilization after case resolution. However, since the final sample size was limited, results should be interpreted as a case study. Despite this limitation, the statistically significant results suggest that online crowdsourcing shows promise as an efficient method of solving difficult medical cases.

Introduction

Patients experiencing complex, non-specific medical symptoms often are misdiagnosed or go undiagnosed for years [1,2]. This problem leads to avoidable health care costs from unnecessary diagnostics and treatments that occur prior to accurate diagnosis, as well as increased patient morbidity and mortality from both unnecessary procedures and treatments, as well as underutilization of appropriate care [3-7]. Many patients visit

multiple physicians seeking an accurate diagnosis, but their access to care is typically limited to their geographic area and health insurance network, as well as by their time availability. Several services seek to expand access to medical care through second-opinion programs, but these programs do not involve interaction with the patient, beyond a review of their medical records, and are limited to one medical expert's opinion [8,9].

With the recent focus on digital innovation in health care, new tools are being developed to facilitate more efficient diagnoses. Various online or mobile phone app-based symptom checkers that use computer algorithms to facilitate self-diagnosis are freely available [10]. However, these tools are intended for fairly typical patient presentations rather than the highly nuanced nature of complex, difficult to diagnose medical cases. Other tools aim to make the second-opinion process more efficient and accessible by bringing it online. One unique new tool specifically focuses on helping patients with difficult medical cases obtain an accurate diagnosis. CrowdMed is an online platform that uses crowdsourcing to overcome the limitations of regional provider access and second-opinion services (CrowdMed, Inc., San Francisco, CA). Crowdsourcing has been proven to be valuable in various fields, for everything from simple to complex tasks, including map generation, logo creation, disease prevalence estimation, and medical image categorization [11-14]. CrowdMed applies the crowdsourcing concept to diagnosis of difficult medical cases by making patient case details available to large groups of people interested in helping solve the case. Patients submit a case containing symptoms, medical history, family history, imaging, and other information regarding their disease. Registered *medical detectives* review the case and suggest potential diagnoses, and a prediction algorithm identifies the most probable diagnosis. More details on CrowdMed can be found on the company website or in Meyer et al. [15].

Meyer et al. previously conducted an independent analysis of the data collected through CrowdMed as an initial assessment of its impact [15]. This assessment was based on self-reported data collected in patient questionnaires as part of the CrowdMed experience. As such, the authors conclude that further empirical validation of the platform's impact is necessary. To analyze its effect on health care utilization, we sought to conduct a more rigorous assessment of the impact of CrowdMed by using medical claims data from patients who completed a case on CrowdMed. Specifically, we explored the hypothesis that using CrowdMed can lower health care utilization by shortening the pathway to an accurate diagnosis. We employed technology-enabled clinical study methods to access and collect retrospective patient-specific medical claims data from a broad range of health insurers across the United States.

Methods

Study Overview

Just as digital innovation has led to new health care tools, it has created opportunities for new study methodologies and processes [16]. In this study, we electronically recruited and enrolled patients who had previously used CrowdMed for a difficult medical diagnosis. We used an online survey to collect information from participants that allowed us to access and retrieve their retrospective medical claims data over a relevant time period through a third-party web application program interface (API). This study was reviewed and approved by Solutions Institutional Review Board (IRB; IRB Identifier 1JUN15-108) and all participants signed informed consent documentation, as well as a Health Insurance Portability and Accountability Act authorization.

Participants and Study Procedures

Patients who completed a case on CrowdMed's platform between July 1, 2014, and April 30, 2015, were at least 18 years of age, and resided in the United States were identified from CrowdMed's user database as eligible for the study. Eligibility was not based on any measure of whether the case was successful or not. Users who had submitted multiple cases were excluded from eligibility. Invitations to the study were sent via email between June 26, 2015, and July 30, 2015, with follow-up phone calls made to bolster enrollment efforts. Users who had previously unsubscribed from the CrowdMed email list did not receive a study invitation, in order to remain compliant with Controlling the Assault of Non-Solicited Pornography and Marketing Act regulations.

Study invitations included a link that brought the patients to the online study interface. Patients first read through information about the study and provided informed eConsent as well as permission to release protected health information for research. Participants were then screened to confirm that they visited at least two different physicians and/or health care facilities in the 12 months prior to using CrowdMed, where the primary reason for the visit was the illness in the CrowdMed case. These parameters were used because CrowdMed is targeted at patients who have already unsuccessfully tried to arrive at a diagnosis for their condition through the traditional health care system.

Study participants were then asked to provide identifying information, and this information was used to access their medical claims data for the 12 months prior to, and time since, their CrowdMed case resolution; a period ranging from 3 to 12 months. This information included member identification numbers and health insurer names for any health insurance plans that participants were enrolled in over that time period, as well as names and addresses for all the physicians and health care facilities they visited in that time period. Entering a national provider identifier (NPI) for each provider was optional. The time periods of 12 months prior to case resolution and 3 to 12 months since resolution were selected to construct a time period (15 to 24 months) which was long enough to capture representative health care utilization, but recent enough that patients could be expected to remember their insurance and provider information over the entire interval.

Claims Data and Measures

Information collected from study participants was used to query a third-party web API in order to access medical claims data. The API is designed to be accessed by automated software and to return information in a consistent, software-readable format, and is hosted by a third-party company that grants access based on negotiated contracts. We wrote software to automatically

query the third-party API for our study participants and quickly pull patient-specific coverage and claims information from different health insurance companies. This approach is more flexible and scalable than being limited to a costly existing claims data set that may or may not include data for patients of interest. The third-party API has wide coverage but does not give access to claims data from every health insurer in the United States, so data was not available for some study participants because of their health insurance coverage. Additionally, the third-party API required us to query it with the name and NPI of the billing provider to access claims data for the study participants; information was not always entered accurately by the study participant, likely because they were unaware of provider billing processes. We cross-referenced provider names and NPIs with others at the same practice or address, in an attempt to find the correct name and NPI for all API calls, but we were not able to pull claims data for all providers entered by participants.

Due to the limitations on data access, study participants were individually assessed for final enrollment eligibility based on the completeness of the set of medical claims accessed, in order to ensure the integrity of the study data. Participants were officially enrolled in the study if they met the following criteria: (1) claims were collected from all insurance carriers under which the respondent had coverage during the relevant time period, and (2) claims were collected from at least 67% of the health care providers entered in the study survey.

Claims data collected via the API included service date(s), Current Procedural Terminology (CPT) codes identifying medical procedures and services received, total charged amount for the claim, total amount paid on the claim by the insurer, and the provider information from the API call. This data was linked with the study participant's CrowdMed data, including demographic information, date of case resolution, and CrowdMed questionnaire results. The combined data set was then stripped of all identifying patient information for analysis, including names, geographic information, telephone numbers, email addresses, medical record numbers, health plan beneficiary numbers, and account numbers.

The primary objective of the study was to assess whether the use of the CrowdMed service decreased health care utilization, which we hypothesized to be the case due to a shortened, more efficient diagnostic pathway with CrowdMed. This endpoint was prespecified as a decrease in monthly frequency of provider visits between the 12 months prior to CrowdMed case resolution and the time since case resolution. Provider visits were defined as one visit per date of service and unique provider (ie, claims from multiple providers on the same day qualify as multiple visits, but claims from the same provider on the same day qualify as one visit). Pharmacy claims were excluded from the visits analysis because they do not involve provider-patient interaction.

Secondary and exploratory analyses examined metrics such as decreases in monthly medical costs, frequency of high-cost medical services or procedures, and complexity of provider visits between pre- and post-CrowdMed case resolution time windows. One measure of complexity of visits was defined

based on CPT codes for evaluation and management (E/M) provider visits. For example, 99213 signifies a *Level 3 Established Patient Office Visit* while code 99214 signifies a *Level 4 Established Patient Office Visit*, where higher levels denote more patient-physician face-to-face time, a more detailed and extensive exam, and more complex medical decision making. We also stratified the utilization results by patient characteristics, such as time since disease onset. For cost analysis, we used charged amount, rather than paid amount, because the amount paid out by the insurer excludes deductibles and co-insurance and therefore underestimates expenditures in a non-uniform pattern. Seasonality is especially problematic with deductibles and biases the pre- and post-time window comparisons. While charged amounts do not accurately reflect expenditures, since insurers actually pay contracted rates for each provider, they are useful for the relative comparisons conducted in this study. We removed duplicate claims for cost analysis, characterizing unique claims based on a unique combination of date of service, provider, charged amount, and CPT codes.

We also performed a comprehensive exploration of data trends over time to rule out other possible temporal confounders. To address questions of whether a period of artificially high health care utilization occurred while the patient had a case actively on CrowdMed (the *medical detectives* often request new diagnostic test results while a case is on the site), and thus biased pre-case resolution utilization as artificially high, we compared utilization metrics from the periods before initiating a case on CrowdMed, while the case was actively on CrowdMed, and post-case resolution. We also investigated whether regression to the mean could influence study results. The hypothesis in this case would be that patients initiate a CrowdMed case at a time when their symptoms are exceptionally severe, and hence the CrowdMed patient pool is biased toward patients experiencing a spike in their symptoms over patients at normal or low symptom levels. After signing up, those patients experiencing a spike in their symptoms would revert to normal symptom levels and would have been expected to improve even without intervention. The signature of this trend would be utilization that spikes right at the time of CrowdMed signup and then tapers off. To test this hypothesis, we examined the directional trend in utilization during the time frame when a case was active on CrowdMed (ie, between case initiation and resolution). A downward trend during this time period would provide evidence that regression to the mean may be present.

Statistical Analysis

We used the one-tailed Wilcoxon signed-rank test to compare data between the pre- and post-CrowdMed case resolution time windows. We chose the Wilcoxon signed-rank test over the paired t-test because its assumptions hold for datasets with small sample sizes and values drawn from skewed distributions. The one-tailed test was used for pre-post comparisons because we were assessing whether the post-resolution utilization was or was not significantly lower than the pre-resolution utilization. When comparing the difference in pre-post utilization change between patient subgroups we used a two-tailed Wilcoxon signed-rank test, since we did not hypothesize a specific direction for a difference. For comparisons between the

characteristics of the study sample and the overall population of CrowdMed users eligible for recruitment, we used the two-tailed versions of the Mann-Whitney *U* test for comparisons of continuous values, and Fisher's exact test for comparisons of proportions.

Statistical tests were performed using R 3.2.2 (Vienna, Austria). Data processing was performed in Python 2.7.10 using the Pandas package [17]. All significance is noted at an alpha level of .05.

Results

Study Sample

Of the 546 CrowdMed patients eligible for recruitment, 445 were sent email invitations to the study, 94 initiated the study survey, and 45 completed it with information for both their health insurer(s) and health care providers (Figure 1). Some amount of medical claims data was accessible via third-party web API for 23 patients. Of these patients, 9 did not meet enrollment eligibility criteria for completeness of their medical claims data set, and 1 was excluded due to irregular claims data of questionable quality. A total of 13 participants were enrolled in the final study sample.

Table 1 presents the characteristics of the study participants, as well as the characteristics of the entire eligible CrowdMed user population. Study participants were directionally more likely to be younger and female, compared to the eligible CrowdMed user population, though neither difference was significant. Study participants were similar to eligible users with regards to ethnicity, time since disease onset, and time since case resolution date. Most study participants had visited fewer than 10 health care providers in the 12 months prior to using CrowdMed.

Table 1. Participant characteristics including standard deviations (SD).

Characteristic	Study Participants (n=13)	CrowdMed Users Eligible for Recruitment (n=546)	Study Participants vs. Eligible Users, *P*-value
Mean age, years (SD)	42.8 (13.5)	51.8 (16.5)	.08[a]
Sex (%)			.28[b]
Female	9 (69%)	289 (52.9%)	
Male	4 (31%)	257 (47.1%)	
Ethnicity (%)			.66[b]
Caucasian/white	11 (85%)	483 (88.5%)	
Other	2 (15%)	63 (11.5%)	
Mean years since disease onset (SD)	6.7 (5.9)	7.6 (10.8)	.46[a]
Mean months between case resolution date and claims data pull (SD)	6.9 (2.8)	8.0 (3.1)	.22[a]
Health care providers visited in the 12 months prior to using CrowdMed, self-reported in screener (%)			N/A
2-4	6 (46%)	N/A	
5-9	6 (46%)	N/A	
10+	1 (8%)	N/A	

[a]*P*-value of a Mann-Whitney *U* test

[b]*P*-value of Fisher's exact test

Figure 1. Participant recruitment and enrollment flowchart. aEligibility criteria: (1) claims were collected from all insurance carriers under which the respondent had coverage during the relevant time period, and (2) claims were collected from at least 67% of the healthcare providers entered in the study survey.
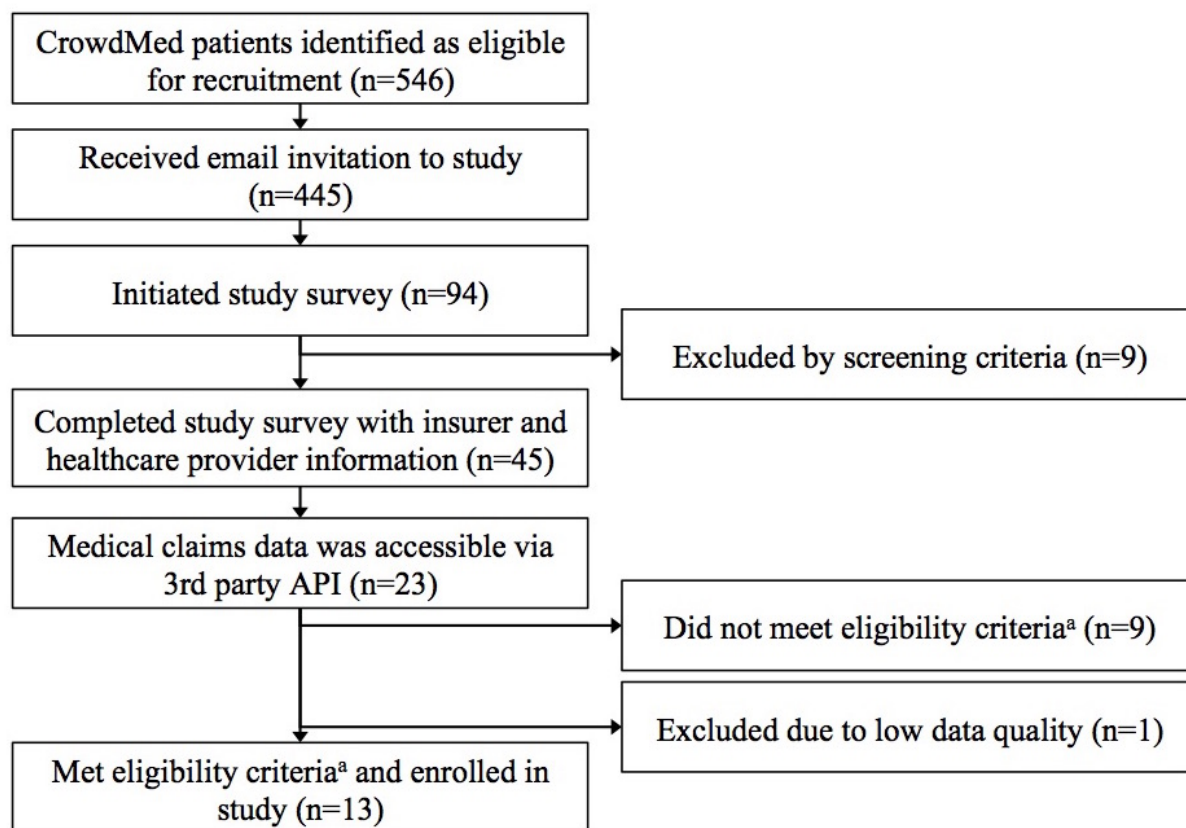
Health Care Utilization

A total of 221 health care provider visits were collected for the study participants (167 prior to CrowdMed case resolution and 54 post-resolution) across 11 US health insurers, with service dates ranging from September 20, 2013, to July 28, 2015. Health care utilization was significantly lower after resolution of a case on CrowdMed compared to before case resolution. The primary endpoint of frequency of health care provider visits decreased significantly, from an average of 1.07 visits per month pre-case resolution to 0.65 visits per month post-case resolution (median visits per month 0.50 vs. 0.34, P=.01, W=79; Figure 2). Medical costs, calculated based on charged amounts as previously described, were also significantly lower after case resolution (mean of US $719.70 per month vs. US $516.79 per month, median of US $349.92 per month vs. US $60.83 per month, P=.03, W=72). Additionally, health care utilization trended downward with time from case resolution date.

We examined metrics and proxies for a decrease in complexity of care in exploratory analyses. The monthly frequency of procedures and services over US $500 was directionally lower after case resolution with a trend toward statistical significance (mean of 0.31 vs. 0.17, median of 0.25 vs. 0.00, P=.06, W=51). Number of CPT codes per visit and average cost per visit were also directionally lower post-case resolution (mean of 3.3 vs. 2.4, median of 3.1 vs. 2.0, P=.19, W=20; and mean of US $470

vs. US $400, median of US $439 vs. US $256, P=.13, W=27), as were the complexity levels of provider visits based on E/M visit CPT codes. Prior to CrowdMed case resolution, 17% of visits were coded as Level 3 office visits (intermediate complexity), while 28% were coded as Level 4 office visits (higher complexity). After CrowdMed case resolution, 24% of provider visits were coded as Level 3 office visits while 19% were coded as Level 4 office visits.

We explored stratifying the study results by various patient characteristics but did not find any factors that significantly correlated with the key result of lower frequency of provider visits post-CrowdMed case resolution. We first examined time since disease onset. Study participants were split by disease onset in 2010 or earlier (n=6) and 2011 or later (n=7). There was no significant difference in the results for change in provider visit frequency post-case resolution between the groups (P=1.00, U=22). We also incorporated a metric that CrowdMed internally uses to designate cases as successful or unsuccessful. This metric is subjective and based on a question that users answer in a survey after their case is resolved, indicating whether the user believes that CrowdMed brought them closer to a correct diagnosis or cure. In our study sample, 9 participants had responded positively while 4 had responded negatively to the self-report question (at 69%, this is similar to the results CrowdMed typically sees across their surveyed user population of 60% to 70%). There was no significant difference between

the two groups in the results for change in provider visit frequency (P=.28, U=11).

In order to assess whether a period of artificially high health care utilization occurs while a case is active on CrowdMed, and biases pre-resolution utilization rates, we split the pre-resolution time period in two: before case initiation versus while active on CrowdMed. We compared utilization during these time periods against each other and against post-case resolution utilization. We found that monthly provider visits were significantly higher during the period that a case was active on CrowdMed compared to before case initiation (mean of 0.99 visits per month before case initiation vs. 1.45 visits per month during CrowdMed use, median of 0.40 vs. 1.00, P=.047, W=70; Figure 3). This result suggests that there is a spike in utilization while a case is active on CrowdMed. To determine if this issue biases study results, we removed the period that a case was active on CrowdMed and compared the before-case-initiation time period with the post-resolution time period. We found that monthly provider visits were significantly lower post-case resolution compared to before case initiation (mean of 0.99

visits per month before case initiation vs. 0.65 visits per month post-case resolution, median of 0.40 vs. 0.34, P=.02, W=65), which is consistent with our base-case results.

Finally, we investigated whether regression to the mean may have influenced study results (ie, patients signed up for CrowdMed when their symptoms were exceptionally severe, and their symptoms would revert to normal levels even in the absence of CrowdMed). This trend would be signified by utilization that peaks at the time of CrowdMed signup and then tapers off. To test this hypothesis, we specifically examined the time frame when a case was active on CrowdMed (between case initiation and resolution) and the directional trend in utilization during this period. A downward trend would provide evidence of regression to the mean. We found a positive correlation of 0.18 between provider visits per week and weeks since CrowdMed case initiation. The correlation between cost per week and weeks since case initiation was similar, at 0.17. This upward trend provides evidence that regression to the mean is not a major determining factor in our study results.

Figure 2. Impact of CrowdMed case resolution on healthcare utilization.
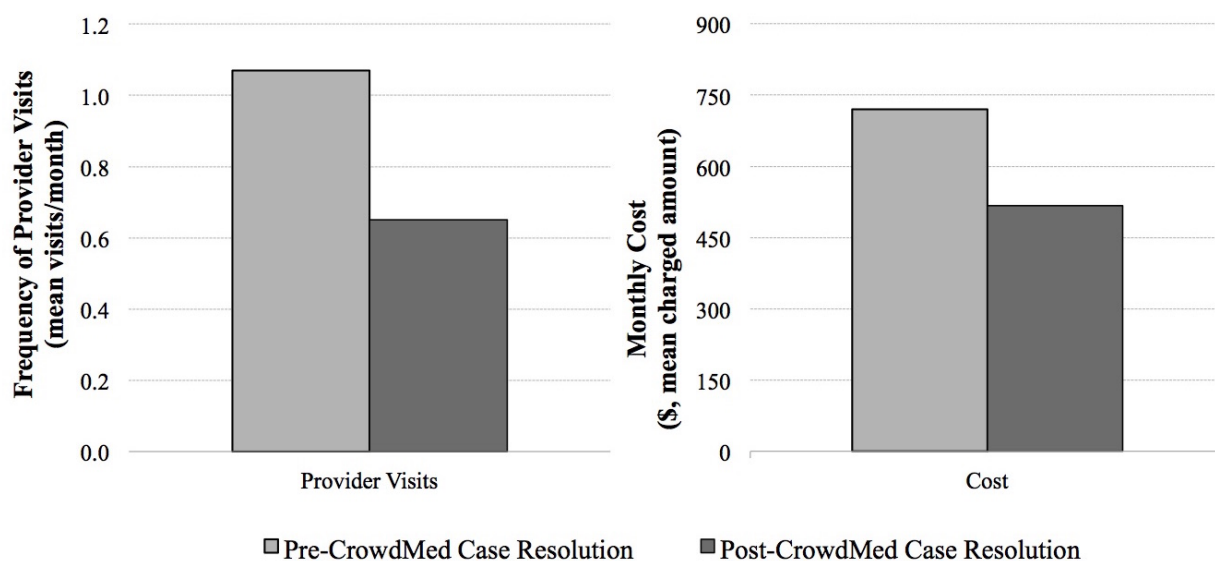
Figure 3. Healthcare utilization before CrowdMed case initiation, during CrowdMed use, and after case resolution.
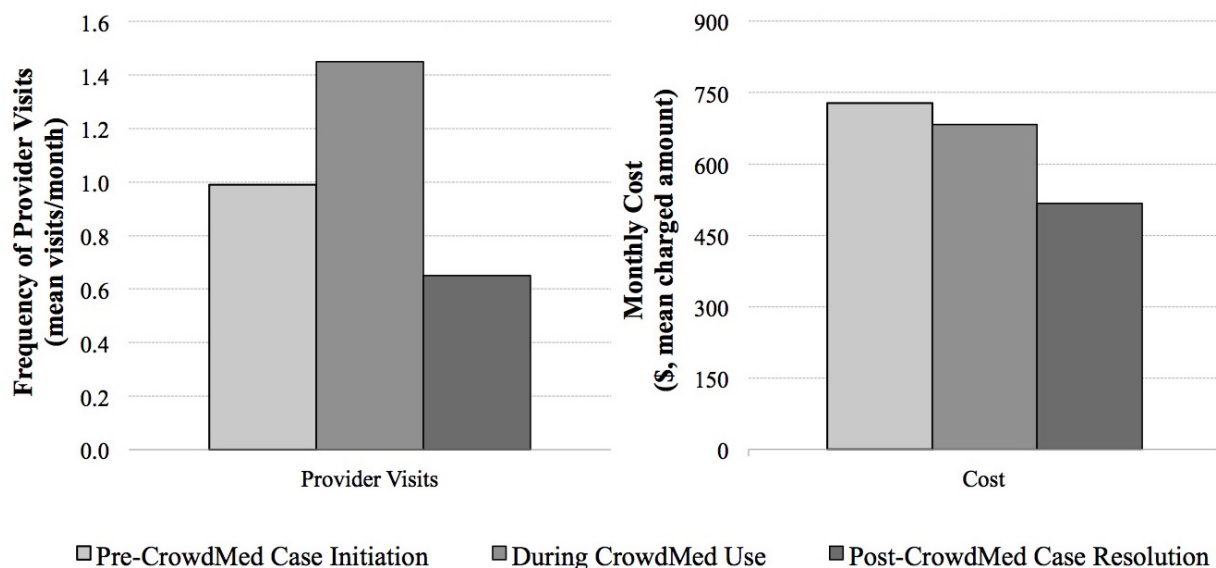
Discussion

Principal Results

This study empirically demonstrated that health care utilization is significantly lower after resolving a difficult medical case on CrowdMed, as evidenced by a small sample of previous CrowdMed users. Both monthly frequency of provider visits and monthly charges were significantly decreased after case resolution per medical claims data. Results were consistent when stratified by various patient characteristics, such as time since disease onset. Additionally, it appears that complexity of care may decrease after case resolution. While results were only directional rather than statistically significant, there were multiple indicators of health care visits and care being less complex after case resolution, such as lower frequency of high-cost procedures, fewer CPT codes per visit, and Level 3 visits becoming more likely than Level 4 visits. The relatively small sample size in the study may explain the lack of statistical significance, rather than a lack of effect.

In addition to the valuable insights that this study provides into the impact of crowdsourcing for diagnosis of difficult medical cases, it also serves as an example of how new technologies can drive innovation in clinical research implementation. Clinical studies have traditionally been limited to site-specific geographies, whereas online recruitment and data collection methods afford broad patient access. Here, we were able to access and recruit CrowdMed users irrespective of where they were located in the United States. Data access for retrospective claims analyses has also been limited in the past to insurer-specific data sets or expensive, often de-identified, data sets for purchase. This is the first study to use a web API to access patient-specific claims data across a broad range of insurers. This new method of claims data access widens the realm of possibilities for claims analyses. Claims data is an

objective indicator of health care utilization, but historically it has been infeasible to collect it from patients themselves. This study demonstrates that new technology enables another effective approach to claims data access, which can be used to test the impact of emerging technologies that may not yet be identifiable through claims codes, or may not have adequate penetration to be studied in any single insurer's separate data set.

The objective nature of a claims analysis is a strength for this study. However, the limitations of claims data may make some aspects of our findings conservative as well as uncertain in a few key manners. For example, claims data excludes quality of life factors. In this study, most participants had lower health care utilization after they resolved their case on CrowdMed. A few, however, had increased provider visit frequency and/or costs. This is not unexpected, as an accurate diagnosis can lead to proper treatment that will improve the patient's quality of life and possibly long-term survival. This treatment could be expensive, but we would need to study the cost-effectiveness of the treatment to determine its overall value. Similarly, CrowdMed could bring a patient closer to an accurate diagnosis by placing that patient on an appropriate diagnostic pathway, which could entail expensive diagnostic tests. Although this may be reflected in our analysis as higher costs post-case resolution, it could represent tests that would be required for an accurate diagnosis, and CrowdMed may have accelerated the steps to those tests. Another conservative aspect of this study (and similar claims analyses) is the exclusion of indirect costs. Provider visits involve cost implications for missed work and travel to appointments, and these costs are borne by patients, caregivers, and employers. While these indirect costs may be irrelevant to some health care payers, they are very impactful to self-insured employers. The ability of CrowdMed to reduce the frequency of provider visits would translate directly to a more productive workforce.

Limitations

This study had several limitations. First, the enrolled sample was small, at 13 patients. As such, we present these results as a case study to acknowledge the potential for lack of generalizability to a larger sample. Despite this, the primary endpoint of change in monthly frequency of provider visits between the pre- and post-case resolution time windows was statistically significant, as was the change in monthly cost. Second, we were limited to charged-amount data for cost analysis, since allowed or contracted amount was not available and insurer paid amount is biased by deductibles. While the relative comparisons we conducted are meaningful with charged-amount data, the absolute numbers themselves are more difficult to interpret. We can apply a modified cost-to-charge ratio that approximates insurer reimbursement rates to provide a more accurate estimate of cost. One estimate is that commercial insurer contracted rates are approximately 70% of charges [18], suggesting that costs are on average US $142 per month lower post-case resolution ($719.70 − $516.79 = $202.91, $202.91 \times 70\% = 142.04). Thirdly, our hypothesis that use of the CrowdMed service decreases health care utilization by shortening the pathway to an accurate diagnosis does not account for the possibility that some patients give up on the diagnostic search and thus underutilize care. For these patients, CrowdMed may be able to improve their care by increasing their utilization. This issue would be a valuable exploration for a subsequent study, as it is outside the scope of this analysis and would require data sources other than claims. Additionally, our current claims-based approach does not discriminate between health care utilization related to the illness in the CrowdMed case or for unrelated reasons. However, there is no reason that unrelated health care utilization should bias our results in one direction over another.

The pre-post study design was also limited by the lack of an external control arm. This issue prompts questions of whether results were biased by an artificial increase in health care utilization during the time a case was active on CrowdMed, or whether regression to the mean influenced results. Our analyses indicate that there is a spike in utilization while a case is active on CrowdMed. This is not surprising, because the CrowdMed *medical detectives* may ask for specific test results to guide their diagnostic suggestions. Despite this increased utilization during the active CrowdMed period, we found that monthly provider visits were still significantly lower post-case resolution compared to before case initiation, which demonstrates that study results hold when controlling for increased utilization during the active case period. As for the possibility that patients initiate a CrowdMed case at an especially intensive period of utilization and disease severity, and would be expected to improve even without intervention, our analyses suggest this is not a factor in our results. Time since disease onset varies in the sample, and results are independent of this patient characteristic, which suggests that patients enter CrowdMed at various points in their disease progression. Additionally, we found that there is an upward trend in frequency of provider visits and cost as a function of time since CrowdMed case initiation during the period the case is active, rather than the downward trend that would be expected if regression to the mean were present.

Conclusions

In this case study, patients who resolved a difficult medical case on the CrowdMed platform were found to have approximately 40% lower frequency of provider visits and 30% lower medical charges after case resolution. This early evidence suggests that wider spread use of crowdsourcing diagnoses in difficult medical cases may have the potential to reduce the burden on the US health care system and lead to more efficient delivery of health care resources. These findings can serve as an initial value assessment for potential purchasers of CrowdMed services and can serve to inform subsequent larger, prospective, controlled studies that bolster the evidence of CrowdMed's impact.

Conflicts of Interest

Joseph Ladapo is a consultant to Evidation Health, Inc., a company specialized in demonstrating digital health product value. Evidation Health, Inc. was paid by CrowdMed, Inc. to conduct this study.

References

1.   Ball JR, Balogh E. Improving diagnosis in health care: highlights of a report from the national academies of sciences, engineering, and medicine. Ann Intern Med 2016 Jan 5;164(1):59-61. [doi: 10.7326/M15-2256] [Medline: 26414299]

XSL•FO

RenderX

2. Schiff G, Kim S, Abrams R, Cosby K, Lambert B, Elstein A, et al. Advances in patient safety: from research to implementation (volume 2: concepts and methodology). Rockville, MD: Agency for Healthcare Research and Quality (US); 2005. Diagnosing diagnosis errors: lessons from a multi-institutional collaborative project URL: http://www.ncbi.nlm.nih.gov/books/NBK20492/pdf/Bookshelf_NBK20492.pdf [accessed 2016-02-09] [WebCite Cache ID 6hGLK2wXo]

3. Schwartz A, Weiner SJ, Weaver F, Yudkowsky R, Sharma G, Binns-Calvey A, et al. Uncharted territory: measuring costs of diagnostic errors outside the medical record. BMJ Qual Saf 2012 Nov;21(11):918-924. [doi: 10.1136/bmjqs-2012-000832] [Medline: 22773889]

4. Shojania KG, Burton EC, McDonald KM, Goldman L. Changes in rates of autopsy-detected diagnostic errors over time: a systematic review. JAMA 2003 Jun 4;289(21):2849-2856. [doi: 10.1001/jama.289.21.2849] [Medline: 12783916]

5. Moore LJ, Jones SL, Kreiner LA, McKinley B, Sucher JF, Todd SR, et al. Validation of a screening tool for the early identification of sepsis. J Trauma 2009 Jun;66(6):1539-46; discussion 1546. [doi: 10.1097/TA.0b013e3181a3ac4b] [Medline: 19509612]

6. Ely JW, Graber ML, Croskerry P. Checklists to reduce diagnostic errors. Acad Med 2011 Mar;86(3):307-313. [doi: 10.1097/ACM.0b013e31820824cd] [Medline: 21248608]

7. McDonald KM, Matesic B, Contopoulos-Ioannidis DG, Lonhart J, Schmidt E, Pineda N, et al. Patient safety strategies targeted at diagnostic errors: a systematic review. Ann Intern Med 2013 Mar 5;158(5 Pt 2):381-389. [doi: 10.7326/0003-4819-158-5-201303051-00004] [Medline: 23460094]

8. Meyer AN, Singh H, Graber ML. Evaluation of outcomes from a national patient-initiated second-opinion program. Am J Med 2015 Oct;128(10):1138.e25-1138.e33. [doi: 10.1016/j.amjmed.2015.04.020] [Medline: 25913850]

9. Reddy ST. The Wall Street Journal. 2015 Aug 24. New ways for patients to get a second opinion: online services from established medical centers and independent businesses URL: http://www.wsj.com/articles/new-ways-to-get-a-second-opinion-1440437584?cb=logged0.4775853557512164 [accessed 2016-02-09] [WebCite Cache ID 6fB9a4bb9]

10. Semigran HL, Linder JA, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. BMJ 2015;351:h3480 [FREE Full text] [Medline: 26157077]

11. Schenk E, Guittard C. Towards a characterization of crowdsourcing practices. J Innov Econ Manage 2011;7(1):93-107.

12. Ranard BL, Ha YP, Meisel ZF, Asch DA, Hill SS, Becker LB, et al. Crowdsourcing--harnessing the masses to advance health and medicine, a systematic review. J Gen Intern Med 2014 Jan;29(1):187-203 [FREE Full text] [doi: 10.1007/s11606-013-2536-8] [Medline: 23843021]

13. Brady CJ, Villanti AC, Pearson JL, Kirchner TR, Gupta OP, Shah CP. Rapid grading of fundus photographs for diabetic retinopathy using crowdsourcing. J Med Internet Res 2014;16(10):e233 [FREE Full text] [doi: 10.2196/jmir.3807] [Medline: 25356929]

14. Mavandadi S, Dimitrov S, Feng S, Yu F, Sikora U, Yaglidere O, et al. Distributed medical image analysis and diagnosis through crowd-sourced games: a malaria case study. PLoS One 2012;7(5):e37245 [FREE Full text] [doi: 10.1371/journal.pone.0037245] [Medline: 22606353]

15. Meyer AN, Longhurst CA, Singh H. Crowdsourcing diagnosis for patients with undiagnosed illnesses: an evaluation of CrowdMed. J Med Internet Res 2016;18(1):e12 [FREE Full text] [doi: 10.2196/jmir.4887] [Medline: 26769236]

16. Rosa C, Campbell AN, Miele GM, Brunner M, Winstanley EL. Using e-technologies in clinical trials. Contemp Clin Trials 2015 Nov;45(Pt A):41-54. [doi: 10.1016/j.cct.2015.07.007] [Medline: 26176884]

17. McKinney W. Data structures for statistical computing in python. 2010 Jun 28 Presented at: Proceedings of the 9th Python in Science Conference; June 28 - July 3 2010; Austin, TX.

18. Hospital charges, costs and payments: understanding terminology.: Oregon Health Authority; 2014 Mar. URL: http://www.oregon.gov/oha/OHPR/RSCH/docs/Hospital_Financials/Hospital_Charges_Prices.pdf [accessed 2016-02-09] [WebCite Cache ID 6fBArase6]

Abbreviations

API: application program interface
CPT: Current Procedural Terminology
E/M: evaluation and management
IRB: Institutional Review Board
NPI: national provider identifier
SD: standard deviation

XSL•FO
RenderX

XSL·FO
RenderX

XSL•FO

RenderX