Journal of Medical Internet Research

Journal Impact Factor (JIF) (2023): 5.8
Volume 20 (2018), Issue 1   ISSN 1438-8871   Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Viewpoints

Original Papers

Reviews

Viewpoint

Facebook Groups for the Management of Chronic Diseases

Stephanie R Partridge[1], PhD, APD; Patrick Gallagher[2], LLB; Becky Freeman[1], PhD; Robyn Gallagher[2], PhD, MN, RN, BA, FAHA, FESC

[1]Prevention Research Collaboration, Sydney School of Public Health, Charles Perkins Centre, The University of Sydney, Sydney, Australia
[2]Sydney Nursing School, Charles Perkins Centre, The University of Sydney, Sydney, Australia

Corresponding Author:
Robyn Gallagher, PhD, MN, RN, BA, FAHA, FESC
Sydney Nursing School
Charles Perkins Centre
The University of Sydney
Room 2210, Level 2, Building D17
The University of Sydney
Sydney, 2006
Australia
Phone: 61 2 86270279
Email: robyn.gallagher@sydney.edu.au

Abstract

The use of Facebook groups by health care researchers and professionals for chronic disease management, namely type 2 diabetes mellitus and coronary heart disease, is in its early stages and challenges are emerging. While Facebook groups offer great potential to deliver health support, research of Facebook groups for chronic disease management remains in its infancy, with robust evidence not yet available. Designing Facebook groups that are acceptable to users, health care researchers as well as health care professionals is a challenge, and there is a poor fit with traditional research and evaluation methods. Key recommendations for future research of Facebook groups for chronic disease management include: (1) iterative content development with input from the target patient population; (2) further understanding of the potential role of group "champions"; (3) ensuring the social media policies of health care institutions allow for real time online communication; and (4) utilizing comprehensive evaluation strategies, including the use of process evaluations.

(J Med Internet Res 2018;20(1):e21) doi:10.2196/jmir.7558

KEYWORDS

social media, prevention, intervention, Facebook

Introduction

Patient education is a core component of chronic disease self-management, and is particularly important for type 2 diabetes mellitus (T2DM) [1] and coronary heart disease (CHD) care [2]. Group-based education provides opportunities for the delivery of detailed information, patient discussions, peer support and direct supervision, and support for behaviors such as exercise [1,2]. There is substantial evidence highlighting the benefits of peer support programs in regard to changing behaviors and reducing risk factors [1,2]. Despite their proven effectiveness, the logistics and costs of staffing and providing the specialized venues for in-person, group-based programs ultimately limits accessibility because services must be offered at fixed and limited times and locations [3]. For instance, attendance at traditional cardiac rehabilitation group-based programs and T2DM group-based self-education is persistently low, at 30% and 48% of those referred, respectively [4,5].

Online, real-time social media platforms, such as Facebook, may offer solutions to existing problems with accessing traditional group-based programs for chronic disease management. In mid-2017, Facebook's community reached two billion people [6]. The continued rise in users is partly due to the growing number of older adults (>65 years) who are joining the social networking site [7]. Recent data suggests that nearly 90% of older adults who were on Facebook reported using the social network to find and share health information [8]. With over 40% of older adults living with two or more chronic conditions [9], the ubiquity of Facebook in their everyday lives has contributed to the emergence of a potential new era of health care information delivery.

Social media interactions enable individuals to read and post material at any time, and from any location, as a part of their usual routine, substantially eliminating obstacles to participation compared to in-person interactions. While expert health care staff are still required, costs may be reduced through more convenient and effective scheduling. However, the potential for Facebook groups to provide novel methods for delivering group-based health care, and enabling support from health care professionals and peers, is yet to be fully harnessed [10].

At this present stage, systematic reviews and meta-analyses are not warranted as previous research has only investigated existing publicly available Facebook groups for general chronic disease management [11], and specifically for T2DM management [12,13], diabetic foot care [14] and hypertension [15]. Encouragingly, two studies are underway investigating the effectiveness of Facebook groups for T2DM [16] and CHD [17]. Other studies with social media groups or features were in young populations or evaluated as a part of larger multicomponent mobile health (mHealth) programs where individual effectiveness of the group could not be determined. Therefore, in this viewpoint, we discuss the issues and potential benefits of using Facebook groups for the management of chronic diseases, namely T2DM and CHD, and provide recommendations for researchers working in this space.

Facebook Groups for the Management of Chronic Diseases

There is emerging evidence that chronic disease self-management programs delivered by alternative means, such as telehealth and electronic health (eHealth) delivery, have comparable outcomes to in-person programs [18,19]. However, recent reviews have attempted to determine the effectiveness of social media, using evidence arising largely from multicomponent telehealth and eHealth interventions, which includes social media features [20-22]. The difficulty with this lack of demarcation is that social media interventions may be more complex than previously thought and create difficulty for replication and implementation into practice [23,24].

Facebook groups may offer a mutual platform of support for the management of chronic diseases. Facebook defines their group feature as “a space to communicate about shared interests with certain people” [25]. The Facebook group feature allows patients and/or health professionals to interact through posts, which includes writing and responding to posts, in a self-subscribing forum [25]. Groups can either be open, closed or private. Closed or private groups are commonly chosen for health research, as only group members can view the content [25]. Other group features include the capacity to allocate moderating privileges to selected members [25].

Facebook groups can enable health care professionals to give both individuals and groups support, advice, and encouragement to foster self-management and behavior change [26,27]. Further health benefits can result through the development of collective knowledge, social networking, and peer-to-peer information exchange. However, to provide robust evidence for replicatation, there are key issues in the development, implementation and evaluation of Facebook groups that require further research.

Development, Implementation and Evaluation of Facebook Groups

Developing Facebook groups that are acceptable to and effective for people with chronic disease, as well as health care professionals and researchers, is a challenge requiring engagement of multidisciplinary teams [28]. While there is good evidence about how best to run in-person peer support groups across a variety of health conditions [29], there is limited guidance for how to develop content and effective engagement and communication strategies for Facebook groups to assist people with the management of T2DM or CHD.

Content for in-person chronic disease management peer support groups cannot be directly converted to group-based interventions on Facebook, due to the difference in communication mechanisms. Pagoto et al [30] developed a model for the adaptation of behavioral interventions for social media delivery. This model provides guidance for content conversion and recommends that the content library aligns with how potential users interact with the Facebook platform [30]. An iterative content design process with input from the target audience has the potential to increase appeal and effectiveness of a Facebook group [31-33].

Formative research on older adults with chronic diseases is lacking, with only two studies investigating cardiac patients’ frequency of social media use [34], and experience and perceptions of using Facebook [35]. One study assessed how patients with T2DM communicate health information using Facebook [36]. Research investigating Facebook groups has been predominantly focused on younger people targeting single behaviors [37-54]. Evidence from such interventions is not generalizable to older populations, considering the differences in Facebook use and communication preferences between the two generations [32].

Moderators of Facebook groups need to be aware of the use and communication preference of older adults living with chronic diseases, as well as being experienced Facebook users themselves [30]. Considering the initial complexity of managing T2DM and CHD, initial group moderation by a health professional may be most appropriate. In addition to the moderators, there is a need to have an existing support network prior to participant enrolment, to avoid the “empty room” phenomenon [28]. This can theoretically be achieved by enrolling peer “champions,” whose role is to actively encourage participants to engage with each other [55]. The role and training of these peer “champions” requires greater understanding, as well as the ideal size of a Facebook support group for a chronic disease management.

Moderators may also be required to provide initial education and reminders to inform group members about the privacy settings of the group, as well as their personal account. All posts within a closed or private Facebook group are only visible to moderators and group members. However, issues such as data security and privacy of data management on commercial platforms, like Facebook, requires further attention.

Health care institutions' policies on the use of social media by health care staff need to be flexible to account for the real-time nature of conversations on Facebook [30]. This is not currently standard practice in many health care institutional policies. For example, some health care institutions require staff, who are representing the institution on Facebook, to have all posts approved by a supervisor one month prior to posting [56]. This hinders not only the continuous and dynamic nature of conversations on Facebook, but also the progression of research in this space.

Analysis of publicly available Facebook groups on chronic disease showed that the majority of groups identified were focused on awareness creation [11,15]. However, the few support groups for patients with chronic disease provided insights about effective content and communication strategies. For example, in the case of T2DM education, Facebook group participation has demonstrated improved knowledge, skills, confidence, and notably improved patient self-management [12]. Higher levels of interaction were seen on posts about peers' personal experiences and realistic self-depictions of living with a chronic disease [13,14]. The usefulness of some groups was associated with the types of posts, and no association was found with the number of likes or presence of user comments [14]. This shows the potential capacity of Facebook groups to offer indirect support [27], and highlights that engagement cannot always be determined by common Facebook analytics, such as the number of likes and comments per post.

The additional challenge of evaluating Facebook groups is that many types of data, including both quantitative and qualitative, must be collected to assess engagement and effectiveness. If the Facebook group is a component of a multicomponent program, process evaluations can potentially provide insight into the causal mechanisms of such interventions and enable fine-grained understanding of the individual components [57,58]. Process evaluation methodology is underutilized in multicomponent intervention research, and this challenges research translation to identify essential intervention components from those that are not as important [59].

Conclusions

No robust evidence presently exists to showcase the advantages and/or disadvantages of using chronic disease peer support groups on Facebook. This is partly because only publicly accessible, peer-led groups have been evaluated or groups have only been evaluated as part of larger multicomponent mHealth programs. While Facebook groups can reduce the participant burden of engaging in in-person group support programs, further research is required to understand their potential future role in chronic disease management.

Acknowledgments

This research was supported by a New South Wales Cardiovascular Research Collaborative Grant Project.

Authors' Contributions

SRP, PG, BF and RG conceptualized the paper. SRP and PG wrote the paper, with input from BF and RG. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Deakin T, McShane CE, Cade JE, Williams RDRR. Group based training for self-management strategies in people with type 2 diabetes mellitus. Cochrane Database Syst Rev 2005 Apr 18(2):CD003417. [doi: 10.1002/14651858.CD003417.pub2] [Medline: 15846663]
2. Woodruffe S, Neubeck L, Clark RA, Gray K, Ferry C, Finan J, et al. Australian Cardiovascular Health and Rehabilitation Association (ACRA) core components of cardiovascular disease secondary prevention and cardiac rehabilitation 2014. Heart Lung Circ 2015 May;24(5):430-441. [doi: 10.1016/j.hlc.2014.12.008] [Medline: 25637253]
3. Clark AM, King-Shier KM, Spaling MA, Duncan AS, Stone JA, Jaglal SB, et al. Factors influencing participation in cardiac rehabilitation programmes after referral and initial attendance: qualitative systematic review and meta-synthesis. Clin Rehabil 2013 Oct;27(10):948-959. [doi: 10.1177/0269215513481046] [Medline: 23798748]
4. Schwennesen N, Henriksen JE, Willaing I. Patient explanations for non-attendance at type 2 diabetes self-management education: a qualitative study. Scand J Caring Sci 2016 Mar;30(1):187-192. [doi: 10.1111/scs.12245] [Medline: 26058576]
5. Neubeck L, Freedman SB, Clark AM, Briffa T, Bauman A, Redfern J. Participating in cardiac rehabilitation: a systematic review and meta-synthesis of qualitative data. Eur J Prev Cardiol 2012 Jun;19(3):494-503. [Medline: 22779092]
6. Facebook Newsroom. 2017 Oct 21. Two Billion People Coming Together on Facebook URL: https://newsroom.fb.com/news/2017/06/two-billion-people-coming-together-on-facebook/ [accessed 2018-01-11] [WebCite Cache ID 6wOek8MDO]
7. Ofcom. Adults's media use and attitudes: report 2017 URL: https://www.ofcom.org.uk/__data/assets/pdf_file/0020/102755/adults-media-use-attitudes-2017.pdf [accessed 2018-01-11] [WebCite Cache ID 6wOenSEs5]

8. Tennant B, Stellefson M, Dodd V, Chaney B, Chaney D, Paige S, et al. eHealth literacy and Web 2.0 health information seeking behaviors among baby boomers and older adults. J Med Internet Res 2015;17(3):e70 [FREE Full text] [doi: 10.2196/jmir.3992] [Medline: 25783036]
9. Australian Institute of Health and Welfare. The burden of chronic respiratory conditions in Australia: a detailed analysis of the Australian Burden of Disease Study 2011. Canberra: AIHW; 2017.
10. Balatsoukas P, Kennedy CM, Buchan I, Powell J, Ainsworth J. The Role of Social Network Technologies in Online Health Promotion: A Narrative Review of Theoretical and Empirical Factors Influencing Intervention Effectiveness. J Med Internet Res 2015;17(6):e141 [FREE Full text] [doi: 10.2196/jmir.3662] [Medline: 26068087]
11. De la Torre-Diez I, Díaz-Pernas FJ, Antón-Rodríguez M. A content analysis of chronic diseases social groups on Facebook and Twitter. Telemed J E Health 2012;18(6):404-408. [doi: 10.1089/tmj.2011.0227] [Medline: 22650380]
12. Zhang Y, He D, Sang Y. Facebook as a platform for health information and communication: a case study of a diabetes group. J Med Syst 2013 Jun;37(3):9942. [doi: 10.1007/s10916-013-9942-7] [Medline: 23588823]
13. Greene JA, Choudhry NK, Kilabuk E, Shrank WH. Online social networking by patients with diabetes: a qualitative evaluation of communication with Facebook. J Gen Intern Med 2011 Mar;26(3):287-292 [FREE Full text] [doi: 10.1007/s11606-010-1526-3] [Medline: 20945113]
14. Abedin T, Al MM, Lasker MAA, Ahmed SW, Shommu N, Rumana N, et al. Social Media as a Platform for Information About Diabetes Foot Care: A Study of Facebook Groups. Can J Diabetes 2017 Feb;41(1):97-101. [doi: 10.1016/j.jcjd.2016.08.217] [Medline: 28126155]
15. Al Mamun M, Ibrahim HM, Turin TC. Social media in communicating health information: an analysis of Facebook groups related to hypertension. Prev Chronic Dis 2015;12:E11 [FREE Full text] [doi: 10.5888/pcd12.140265] [Medline: 25633486]
16. Boudreau F, Moreau M, Côté J. Effectiveness of Computer Tailoring Versus Peer Support Web-Based Interventions in Promoting Physical Activity Among Insufficiently Active Canadian Adults With Type 2 Diabetes: Protocol for a Randomized Controlled Trial. JMIR Res Protoc 2016;5(1):e20 [FREE Full text] [doi: 10.2196/resprot.5019] [Medline: 26869015]
17. Siegmund LA, Ahmed HM, Crawford MT, Bena JF. Feasibility of a Facebook Intervention for Exercise Motivation and Cardiac Rehabilitation Adherence: Study Protocol. JMIR Res Protoc 2017 Aug 18;6(8):e162 [FREE Full text] [doi: 10.2196/resprot.7554] [Medline: 28821473]
18. Neubeck L, Redfern J, Fernandez R, Briffa T, Bauman A, Freedman SB. Telehealth interventions for the secondary prevention of coronary heart disease: a systematic review. Eur J Cardiovasc Prev Rehabil 2009 Jun;16(3):281-289. [doi: 10.1097/HJR.0b013e32832a4e7a] [Medline: 19407659]
19. Lee SWH, Chan CKY, Chua SS, Chaiyakunapruk N. Comparative effectiveness of telemedicine strategies on type 2 diabetes management: A systematic review and network meta-analysis. Sci Rep 2017 Oct 04;7(1):12680. [doi: 10.1038/s41598-017-12987-z] [Medline: 28978949]
20. 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]
21. Williams G, Hamm MP, Shulhan J, Vandermeer B, Hartling L. Social media interventions for diet and exercise behaviours: a systematic review and meta-analysis of randomised controlled trials. BMJ Open 2014;4(2):e003926 [FREE Full text] [doi: 10.1136/bmjopen-2013-003926] [Medline: 24525388]
22. Merolli M, Gray K, Martin-Sanchez F. Health outcomes and related effects of using social media in chronic disease management: a literature review and analysis of affordances. J Biomed Inform 2013 Dec;46(6):957-969 [FREE Full text] [doi: 10.1016/j.jbi.2013.04.010] [Medline: 23702104]
23. Petticrew M, Anderson L, Elder R, Grimshaw J, Hopkins D, Hahn R, et al. Complex interventions and their implications for systematic reviews: a pragmatic approach. J Clin Epidemiol 2013 Nov;66(11):1209-1214. [doi: 10.1016/j.jclinepi.2013.06.004] [Medline: 23953085]
24. Petticrew M. Time to rethink the systematic review catechism? Moving from 'what works' to 'what happens'. Syst Rev 2015 Mar 28;4:36 [FREE Full text] [doi: 10.1186/s13643-015-0027-1] [Medline: 25875303]
25. Facebook. Group Basics. 2016. URL: https://www.facebook.com/unsupportedbrowser [accessed 2018-01-11] [WebCite Cache ID 6wOf4ciJg]
26. 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]
27. 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]
28. Cobb NK, Graham AL. Health behavior interventions in the age of facebook. Am J Prev Med 2012 Nov;43(5):571-572. [doi: 10.1016/j.amepre.2012.08.001] [Medline: 23079184]
29. Embuldeniya G, Veinot P, Bell E, Bell M, Nyhof-Young J, Sale J, et al. The experience and impact of chronic disease peer support interventions: a qualitative synthesis. Patient Educ Couns 2013 Jul;92(1):3-12 [FREE Full text] [doi: 10.1016/j.pec.2013.02.002] [Medline: 23453850]

30. Pagoto S, Waring ME, May CN, Ding EY, Kunz WH, Hayes R, et al. Adapting Behavioral Interventions for Social Media Delivery. J Med Internet Res 2016;18(1):e24 [FREE Full text] [doi: 10.2196/jmir.5086] [Medline: 26825969]
31. Xie B, Watkins I, Golbeck J, Huang M. Understanding and Changing Older Adults' Perceptions and Learning of Social Media. Educ Gerontol 2012 Apr 01;38(4):282-296 [FREE Full text] [doi: 10.1080/03601277.2010.544580] [Medline: 22639483]
32. Leist AK. Social media use of older adults: a mini-review. Gerontology 2013;59(4):378-384 [FREE Full text] [doi: 10.1159/000346818] [Medline: 23594915]
33. Schein R, Wilson KK. Literature review on effectiveness of the use of social media: a report for Peel public health. 2010. URL: https://www.peelregion.ca/health/resources/pdf/socialmedia.pdf [accessed 2018-01-11] [WebCite Cache ID 6wOfLb1tm]
34. Gallagher R, Roach K, Sadler L, Glinatsis H, Belshaw J, Kirkness A, et al. Mobile Technology Use Across Age Groups in Patients Eligible for Cardiac Rehabilitation: Survey Study. JMIR Mhealth Uhealth 2017 Oct 24;5(10):e161. [doi: 10.2196/mhealth.8352]
35. Partridge SR, Grunseit AC, Gallagher P, Freeman B, O'Hara BJ, Neubeck L, et al. Cardiac Patients' Experiences and Perceptions of Social Media: Mixed-Methods Study. J Med Internet Res 2017 Sep 15;19(9):e323 [FREE Full text] [doi: 10.2196/jmir.8081] [Medline: 28916507]
36. Menefee HK, Thompson MJ, Guterbock TM, Williams IC, Valdez RS. Mechanisms of Communicating Health Information Through Facebook: Implications for Consumer Health Information Technology Design. J Med Internet Res 2016 Aug 11;18(8):e218 [FREE Full text] [doi: 10.2196/jmir.5949] [Medline: 27515151]
37. Joseph RP, Keller C, Adams MA, Ainsworth BE. Print versus a culturally-relevant Facebook and text message delivered intervention to promote physical activity in African American women: a randomized pilot trial. BMC Womens Health 2015 Mar 27;15:30 [FREE Full text] [doi: 10.1186/s12905-015-0186-1] [Medline: 25886945]
38. 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]
39. Bull SS, Levine DK, Black SR, Schmiege SJ, Santelli J. Social media-delivered sexual health intervention: a cluster randomized controlled trial. Am J Prev Med 2012 Nov;43(5):467-474 [FREE Full text] [doi: 10.1016/j.amepre.2012.07.022] [Medline: 23079168]
40. Patrick K, Marshall SJ, Davila EP, Kolodziejczyk JK, Fowler JH, Calfas KJ, et al. Design and implementation of a randomized controlled social and mobile weight loss trial for young adults (project SMART). Contemp Clin Trials 2014 Jan;37(1):10-18 [FREE Full text] [doi: 10.1016/j.cct.2013.11.001] [Medline: 24215774]
41. 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]
42. 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]
43. 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]
44. Wojcicki JM, Geissler JD, Stokes CW, Heyman MB, Tran CT. The use of the RESPeRATE device to lower blood pressure in inner city obese adolescents and children: a pilot feasibility study. High Blood Press Cardiovasc Prev 2013 Jun;20(2):89-92 [FREE Full text] [doi: 10.1007/s40292-013-0014-3] [Medline: 23653175]
45. Herring SJ, Cruice JF, Bennett GG, Davey A, Foster GD. Using technology to promote postpartum weight loss in urban, low-income mothers: a pilot randomized controlled trial. J Nutr Educ Behav 2014;46(6):610-615. [doi: 10.1016/j.jneb.2014.06.002] [Medline: 25069621]
46. Young SD, Cumberland WG, Lee S, Jaganath D, Szekeres G, Coates T. Social networking technologies as an emerging tool for HIV prevention: a cluster randomized trial. Ann Intern Med 2013 Sep 3;159(5):318-324 [FREE Full text] [doi: 10.7326/0003-4819-159-5-201309030-00005] [Medline: 24026317]
47. 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]
48. Struik LL, Baskerville NB. The role of Facebook in Crush the Crave, a mobile- and social media-based smoking cessation intervention: qualitative framework analysis of posts. J Med Internet Res 2014;16(7):e170 [FREE Full text] [doi: 10.2196/jmir.3189] [Medline: 25016998]
49. Lelutiu-Weinberger C, Gamarel KE, Golub SA, Parsons JT. Race-based differentials in the impact of mental health and stigma on HIV risk among young men who have sex with men. Health Psychol 2015 Aug;34(8):847-856. [doi: 10.1037/hea0000192] [Medline: 25545041]

50. Rice E, Tulbert E, Cederbaum J, Barman Adhikari A, Milburn NG. Mobilizing homeless youth for HIV prevention: a social network analysis of the acceptability of a face-to-face and online social networking intervention. Health Education Research 2012 Jan 13;27(2):226-236. [doi: 10.1093/her/cyr113]
51. Ridout B, Campbell A. Using Facebook to deliver a social norm intervention to reduce problem drinking at university. Drug Alcohol Rev 2014 Nov;33(6):667-673. [doi: 10.1111/dar.12141] [Medline: 24689339]
52. 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]
53. 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]
54. Kim SJ, Marsch LA, Brunette MF, Dallery J. Harnessing Facebook for Smoking Reduction and Cessation Interventions: Facebook User Engagement and Social Support Predict Smoking Reduction. J Med Internet Res 2017 May 23;19(5):e168 [FREE Full text] [doi: 10.2196/jmir.6681] [Medline: 28536096]
55. Fleury J, Keller C, Perez A, Lee SM. The role of lay health advisors in cardiovascular risk reduction: a review. Am J Community Psychol 2009 Sep;44(1-2):28-42. [doi: 10.1007/s10464-009-9253-9] [Medline: 19533327]
56. South Eastern Sydney Local Health District. 2015. Business Rule - SESLHD Use of Facebook NSW Government URL: http://www.seslhd.health.nsw.gov.au/Policies_Procedures_Guidelines/Corporate/Communications/Documents/SESLHDUseofFacebook.pdf [accessed 2018-01-11] [WebCite Cache ID 6wOfVCWYJ]
57. Liu H, Muhunthan J, Hayek A, Hackett M, Laba T, Peiris D, et al. Examining the use of process evaluations of randomised controlled trials of complex interventions addressing chronic disease in primary health care-a systematic review protocol. Syst Rev 2016 Aug 15;5(1):138 [FREE Full text] [doi: 10.1186/s13643-016-0314-5] [Medline: 27526851]
58. 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]
59. Guise J, Chang C, Viswanathan M, Glick S, Treadwell J, Umscheid C. Systematic Reviews of Complex Multicomponent Health Care Interventions Rockville (MD): Agency for Healthcare Research and Quality (US). 2014. URL: https://www.ncbi.nlm.nih.gov/books/NBK194846/ [WebCite Cache ID 6wOfeSdrk]

Abbreviations

CHD: coronary heart disease
eHealth: electronic health
mHealth: mobile health
T2DM: type 2 diabetes mellitus

Edited by G Eysenbach; submitted 27.02.17; peer-reviewed by P Balatsoukas, TR Soron, M Ashford, PDJ Ehlers, A Cyr, J Rumbold, T Bose; comments to author 12.07.17; revised version received 07.11.17; accepted 04.12.17; published 17.01.18.

Please cite as:
Partridge SR, Gallagher P, Freeman B, Gallagher R
Facebook Groups for the Management of Chronic Diseases
J Med Internet Res 2018;20(1):e21
URL: http://www.jmir.org/2018/1/e21/
doi:10.2196/jmir.7558
PMID:29343460

Original Paper

A Perioperative eHealth Program to Enhance Postoperative Recovery After Abdominal Surgery: Process Evaluation of a Randomized Controlled Trial

Eva van der Meij[1,2], MD; Judith AF Huirne[1,2], MD, PhD; A Dorien ten Cate[3], MD; Hein BAC Stockmann[4], MD, PhD; Piet C Scholten[5], MD, PhD; Paul HP Davids[6], MD, PhD; H Jaap Bonjer[7], MD, PhD; Johannes R Anema[1], MD, PhD

[1]Amsterdam Public Health Research Institute, Department of Public and Occupational Health, VU University Medical Center, Amsterdam, Netherlands
[2]Department of Obstetrics and Gynaecology, VU University Medical Center, Amsterdam, Netherlands
[3]Department of Obstetrics and Gynaecology, Spaarne Gasthuis, Haarlem, Netherlands
[4]Department of Surgery, Spaarne Gasthuis, Haarlem, Netherlands
[5]Department of Obstetrics and Gynaecology, Diakonessenhuis, Utrecht, Netherlands
[6]Department of Surgery, Diakonessenhuis, Utrecht, Netherlands
[7]Department of Surgery, VU University Medical Center, Amsterdam, Netherlands

Corresponding Author:
Eva van der Meij, MD
Amsterdam Public Health Research Institute
Department of Public and Occupational Health
VU University Medical Center
van der Boechorststraat 7
1081 BT
Amsterdam,
Netherlands
Phone: 31 204450703
Email: ev.vandermeij@vumc.nl

Abstract

Background: Electronic health (eHealth) interventions have proven effective, but implementation in clinical practice is difficult. More research focusing on the implementation process of eHealth interventions is necessary.

Objective: The objective of this study was to describe the process evaluation of a perioperative eHealth intervention, aiming to enhance recovery after laparoscopic abdominal surgery.

Methods: A process evaluation was carried out alongside a multicenter randomized controlled trial. Patients aged between 18 and 75 years who were scheduled for a laparoscopic cholecystectomy, hernia inguinal surgery, or laparoscopic adnexal surgery were included. The eHealth intervention comprised a website and mobile phone app with the possibility to develop a personalized convalescence plan, a section with information about the surgical procedure and the recovery period, the possibility to ask questions via an electronic consultation (eConsult), and an activity tracker. The process evaluation was carried out using the model of Linnan and Steckler, measuring components such as reach, dose delivered, dose received, fidelity, and participants' attitudes. Implementation scores were calculated based on the average of the four components. Quantitative data were collected by means of an electronic questionnaire, a logistic database, a weblog, and medical files. Qualitative data were collected by conducting interviews with a subsample of the study participants.

Results: A total of 344 of the 863 eligible patients were included in the study, which accounted for a reach of 39.9%, and 173 participants were randomized to the intervention group. The implementation scores of the different functions of the intervention ranged between 60% and 65%. The website, mobile phone app, and activity tracker were rated 7.3 to 7.6 on a scale of 1 to 10. Almost all participants who were interviewed about the eConsult function rated it as being of additional value if combined with the usual care but not as a replacement for usual care.

Conclusions: Although participants were overall satisfied with the intervention, the implementation scores of the different functions of the intervention were fair. More research is needed to evaluate the barriers and facilitators for implementation of this perioperative eHealth intervention in normal practice outside study setting.

Trial Registration: Netherlands Trial Registry NTR4699; http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=4699 (Archived by WebCite at http://www.webcitation.org/6vr02V4KK)

(J Med Internet Res 2018;20(1):e1) doi:10.2196/jmir.8338

KEYWORDS

telemedicine; cholecystectomy; surgical procedures, operative; perioperative care; convalescence; process assessment

Introduction

The conviction that patients heal better in their own personal environment has been one of the main drivers reducing hospital stay after surgical interventions. Cutting direct hospitalization costs while increasing revenues due to more efficient use of hospital resources is another strong incentive. However, the transition from in-hospital recovery to domestic convalescence has occurred at a high pace without sufficient attention to the needs of patients [1]. As a consequence, the length of recovery after surgery takes longer than the period considered to be needed from a medical perspective [2-6]. Literature shows that patients deal with feelings of uncertainty regarding their recovery when they are at home, and in addition, it is proven that influencing these feelings by education and support would have a positive effect on the length of recovery [6]. Therefore, an electronic health (eHealth) intervention focusing on information supply and guidance during the perioperative period of commonly applied gynecological surgical procedures was developed [7,8]. The effectivity of the eHealth intervention was evaluated in two different trials; patients who used the eHealth intervention in the perioperative period returned to work earlier, reported higher quality of life scores, and lower pain scores than patients who received usual perioperative care only [7,9]. Therefore, the intervention was further developed; new features, such as a mobile phone app, an activity tracker, and an electronic consultation (eConsult) function, were added. In addition, the intervention was extended, whereby it could also be used in the perioperative period of commonly applied general surgical procedures [10]. Due to these promising results regarding the effectivity of the intervention, implementation of the intervention in clinical practice seems logical. However, although literature shows in general that eHealth interventions can show beneficial effects, execution of these types of interventions in clinical practice has often been slower and more difficult than expected [11,12]. To evaluate whether the eHealth intervention was executed as planned, we conducted a process evaluation. The aim is to systematically analyze the process from offering the different aspects of the intervention to the participant. By doing this, the feasibility of the intervention will be investigated and barriers and facilitators for future implementation could be explored. In addition, evaluating the adherence to the intervention protocol should be an integral part of evaluating this type of interventions, as this will play an important role in interpreting the results regarding the effectivity.

Methods

Trial Design

This process evaluation was carried out alongside a multicenter randomized clinical trial in seven teaching hospitals in the Netherlands (Netherlands Trial Registry NTR4699). A detailed description of the study design has been published earlier in this journal [10]. The study was reported in accordance with Consolidated Standards of Reporting Trials of Electronic and Mobile HEalth Applications and onLine TeleHealth (CONSORT-EHEALTH) [13] and was approved by the local medical ethics committee under the registration number 2014.301.

Participants

Patients aged between 18 and 75 years who were scheduled for a laparoscopic cholecystectomy, hernia inguinal surgery, or laparoscopic adnexal surgery were eligible to participate. A sample size calculation was performed; a total of 308 participants would be required. More details about the study population and sample size calculation have been described in the study protocol [10].

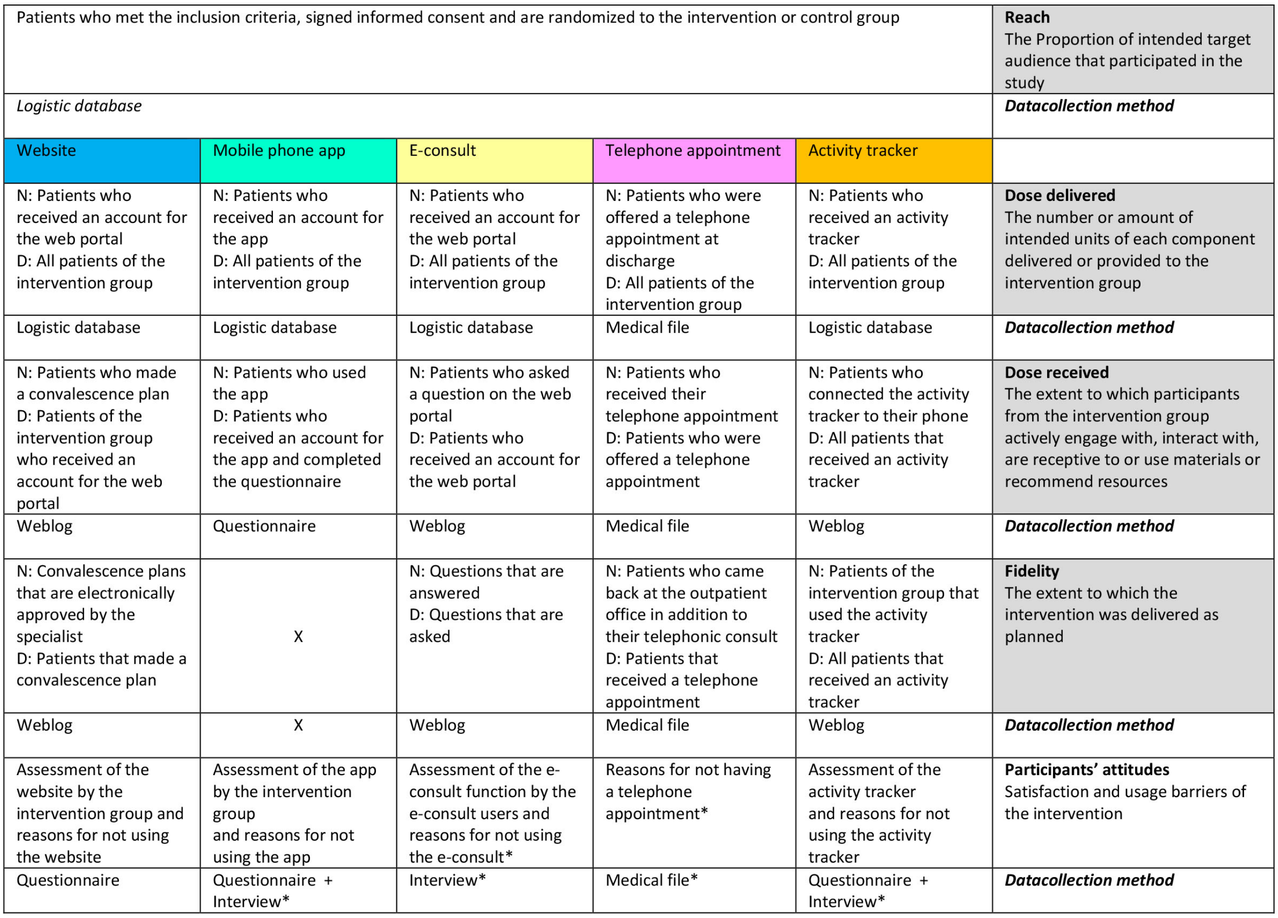
Interventions

Participants were randomized to the control group or the intervention group. Participants from the control group received the usual care and access to a placebo website (containing the patient information brochure about the surgical procedure). Participants from the intervention group received access to the eHealth care program consisting of a website, a mobile phone app, and an activity tracker. The most important tools of the website were the possibility to develop a personalized convalescence plan and the possibility to ask questions to the health care professional (eConsult). Because participants were provided with the option to ask questions via the website by an eConsult, they were initially only offered a telephonic appointment instead of an appointment in the outpatient clinic [10].

Study Settings

Quantitative data were collected 3 months after surgery by an electronic questionnaire, a logistic database, a weblog, and medical files. In addition, qualitative data were collected by conducting telephone interviews. By means of purposive sampling, a sample of participants was selected from the total study population for an additional interview to collect some additional information for this study. The sample consisted of a subsample of participants who used the eConsult function and a subsample of participants who did not. First, a patient who used the eConsult function was selected, and when this patient was willing to be interviewed, a participant who did not use the eConsult function was selected and matched according to age, gender, and type of surgical procedure. This was repeated until data saturation was reached. This approach enables the exploration of the opinion of the participants about the intervention, including reasons for using or not using the eConsult.

Figure 1. Description of the outcome measures . N: nominator; D: denominator. Asterisk indicates data collection methods not described in the original protocol.



Patients who met the inclusion criteria, signed informed consent and are randomized to the intervention or control group					**Reach** The Proportion of intended target audience that participated in the study
Logistic database					***Datacollection method***
Website	Mobile phone app	E-consult	Telephone appointment	Activity tracker	
N: Patients who received an account for the web portal D: All patients of the intervention group	N: Patients who received an account for the app D: All patients of the intervention group	N: Patients who received an account for the web portal D: All patients of the intervention group	N: Patients who were offered a telephone appointment at discharge D: All patients of the intervention group	N: Patients who received an activity tracker D: All patients of the intervention group	**Dose delivered** The number or amount of intended units of each component delivered or provided to the intervention group
Logistic database	Logistic database	Logistic database	Medical file	Logistic database	***Datacollection method***
N: Patients who made a convalescence plan D: Patients of the intervention group who received an account for the web portal	N: Patients who used the app D: Patients who received an account for the app and completed the questionnaire	N: Patients who asked a question on the web portal D: Patients who received an account for the web portal	N: Patients who received their telephone appointment D: Patients who were offered a telephone appointment	N: Patients who connected the activity tracker to their phone D: All patients that received an activity tracker	**Dose received** The extent to which participants from the intervention group actively engage with, interact with, are receptive to or use materials or recommend resources
Weblog	Questionnaire	Weblog	Medical file	Weblog	***Datacollection method***
N: Convalescence plans that are electronically approved by the specialist D: Patients that made a convalescence plan	X	N: Questions that are answered D: Questions that are asked	N: Patients who came back at the outpatient office in addition to their telephonic consult D: Patients that received a telephone appointment	N: Patients of the intervention group that used the activity tracker D: All patients that received an activity tracker	**Fidelity** The extent to which the intervention was delivered as planned
Weblog	X	Weblog	Medical file	Weblog	***Datacollection method***
Assessment of the website by the intervention group and reasons for not using the website	Assessment of the app by the intervention group and reasons for not using the app	Assessment of the e-consult function by the e-consult users and reasons for not using the e-consult*	Reasons for not having a telephone appointment*	Assessment of the activity tracker and reasons for not using the activity tracker	**Participants' attitudes** Satisfaction and usage barriers of the intervention
Questionnaire	Questionnaire + Interview*	Interview*	Medical file*	Questionnaire + Interview*	***Datacollection method***

Outcomes

The process evaluation was performed using the model of Linnan and Steckler [14]. This is a commonly used model in this research field and has the potential to evaluate the process of the implementation systematically because it describes the adherence to the intervention in 5 terms: reach, dose delivered, dose received, fidelity, and participants' attitudes. Except for the reach component of the model, the components were assessed for each function of the intervention separately. The definitions of the different components of the model are presented in Figure 1. A detailed description of the definitions is provided in Multimedia Appendix 1.

Randomization and Blinding

Participants were randomized to the intervention or the control group in a 1:1 ratio by a researcher who was independent from the recruitment, data collection process, or analyses. The study participants were blinded to the allocation.

Data Analysis

IBM SPSS Statistics version 20.0 was used for analyzing the data. The quantitative data were analyzed using descriptive statistics such as frequencies, means, and standard deviations (SD). Implementation scores were calculated using the averaging approach, which means that the average scores of the process measures (reach, dose delivered, dose received, and fidelity) were calculated for each function of the intervention. Qualitative data were transcribed verbatim.

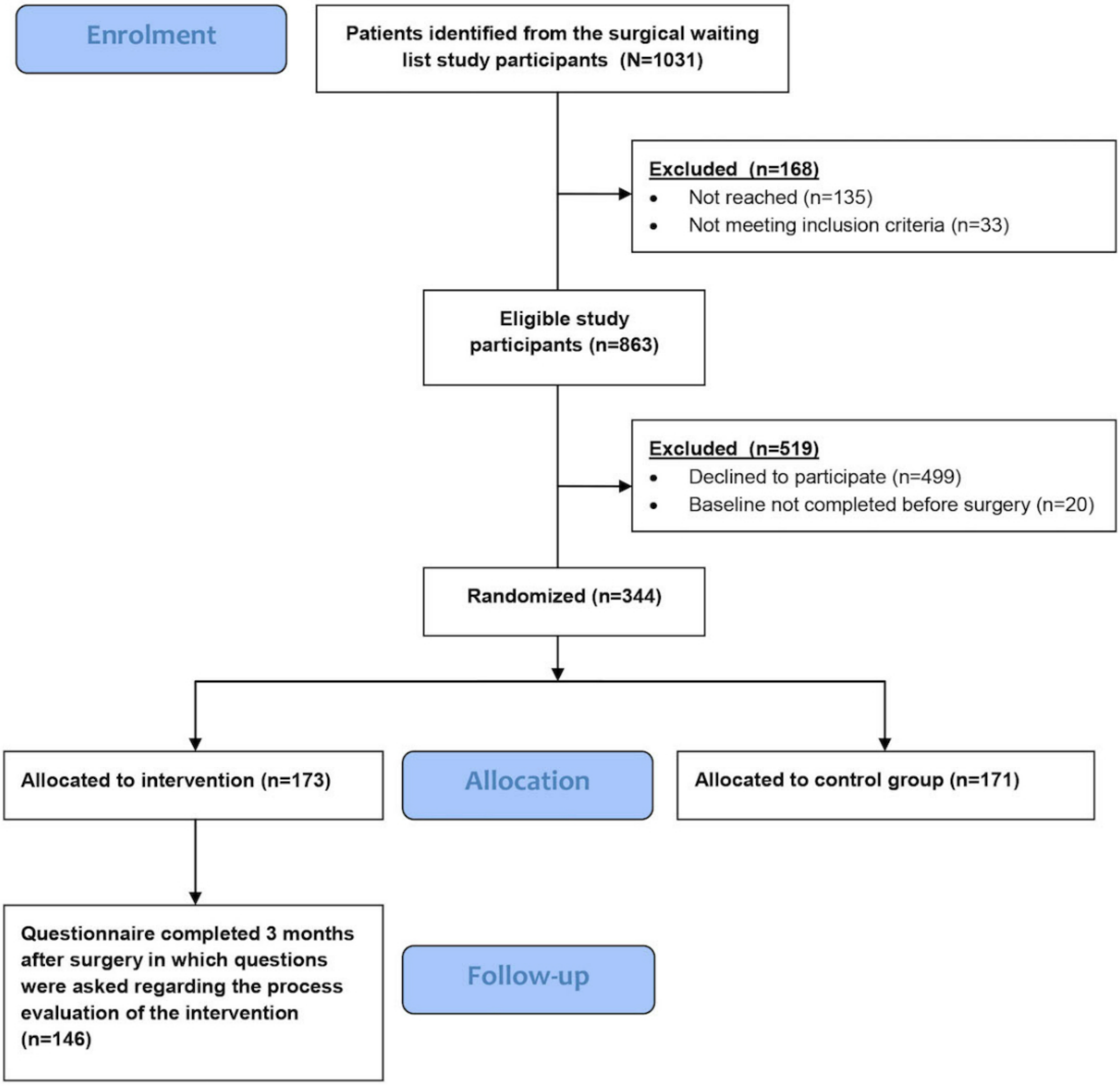
Results

Reach

During September 2015 and August 2016, 1031 potential participants were identified from the surgical waiting list. The flow of the inclusion process has been described in Figure 2. A total of 344 participants gave consent to participate (39.9%, 344/863); there were no major differences regarding age, gender, and surgical procedure between the participants and nonparticipants.

A total of 173 participants were randomized to the intervention group. The baseline characteristics of these participants are presented in Table 1. In addition, 45.1% were male and the mean age was 51 years; 54 participants underwent adnexal surgery, 68 hernia inguinal surgery (1 open procedure), and 51 a cholecystectomy. The response rate to the questionnaire which was assessed 3 months after surgery, in which questions were asked according to participants' attitudes regarding the different functions of the intervention and the usage of the app, was 84.4% (146/173). By purposive sampling, participants were selected for an additional interview. After 12 interviews (6 with participants who used the eConsult function and 6 with participants who did not) data saturation was reached.

Figure 2. Flow diagram. Only the follow-up results concerning the process evaluation are presented.



The subsample of eConsult users consisted of 2 males and 4 females. Three of them underwent a laparoscopic cholecystectomy, 2 hernia inguinal surgery, and 1 adnexal surgery. The mean age was 39 years. Due to the purposive sampling method, the subsample of participants who did not make use of the eConsult function had the same composition regarding gender and surgical procedure. The mean age in this subsample was 44 years. The other components of the Linnan and Steckler model will be described separately for each part of the intervention and are also presented in Figure 3.

Website

Dose Delivered

A total of 172 of the 173 participants of the intervention group (99.4%) received an account for the website. One patient did not receive an account due to logistic problems.

Dose Received

Of the 172 participants in the intervention group who received an account for the website, 138 (80.2%) developed a convalescence plan on the website.

Fidelity

Only 25.2% of the convalescence plans were electronically approved by the medical specialists.

Participants' Attitudes

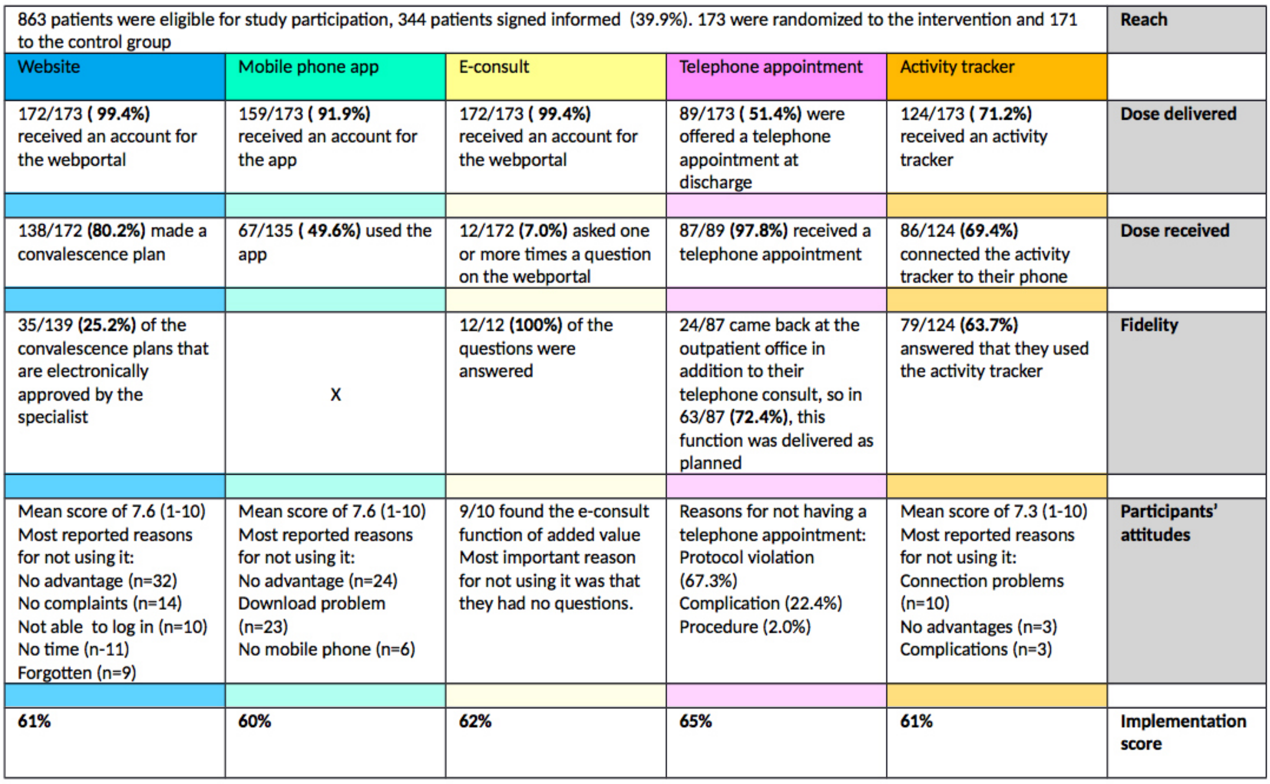
Participants assessed the website with a mean score of 7.6 on a scale of 1 to 10. Reasons for not (frequently) using the website were that participants reported that they did not see the added value (n=32), had no need for it because they had no complaints (n=14), were not able to log in (n=10), had no time (n=11), had forgotten that there was a website (n=9), underwent another type of surgery (open procedure instead of laparoscopic approach) (n=2), used the app (n=2), did not find the information that they were looking for (n=1), or had no computer (n=1).

Table 1. Baseline characteristics.

Variable	Intervention group (N=173)
Gender, n (%)	
Male	78 (45.1)
Female	95 (54.9)
Age (mean, SD[a])	51 (12.57)
Nationality, n (%)	
Dutch	171 (98.8)
Other	2 (1.2)
Level of education, n (%)	
Low	31 (17.9)
Medium	50 (28.9)
High	92 (53.2)
Working situation, n (%)	
Employed	132 (76.3)
Not employed	41 (23.7)
Type of surgery, n (%)	
Laparoscopic adnexal surgery	54 (31.2)
Laparoscopic hernia inguinal surgery	67 (38.7)
Open hernia inguinal surgery	1 (0.6)
Laparoscopic cholecystectomy	51 (29.5)

[a] SD: standard deviation.

Figure 3. Results of the process evaluation.



863 patients were eligible for study participation, 344 patients signed informed (39.9%). 173 were randomized to the intervention and 171 to the control group					**Reach**
Website	Mobile phone app	E-consult	Telephone appointment	Activity tracker	
172/173 (**99.4%)** received an account for the webportal	159/173 (**91.9%)** received an account for the app	172/173 (**99.4%)** received an account for the webportal	89/173 (**51.4%)** were offered a telephone appointment at discharge	124/173 (**71.2%)** received an activity tracker	**Dose delivered**
138/172 **(80.2%)** made a convalescence plan	67/135 (**49.6%)** used the app	12/172 **(7.0%)** asked one or more times a question on the webportal	87/89 **(97.8%)** received a telephone appointment	86/124 **(69.4%)** connected the activity tracker to their phone	**Dose received**
35/139 **(25.2%)** of the convalescence plans that are electronically approved by the specialist	X	12/12 **(100%)** of the questions were answered	24/87 came back at the outpatient office in addition to their telephone consult, so in 63/87 **(72.4%)**, this function was delivered as planned	79/124 **(63.7%)** answered that they used the activity tracker	**Fidelity**
Mean score of 7.6 (1-10) Most reported reasons for not using it: No advantage (n=32) No complaints (n=14) Not able to log in (n=10) No time (n-11) Forgotten (n=9)	Mean score of 7.6 (1-10) Most reported reasons for not using it: No advantage (n=24) Download problem (n=23) No mobile phone (n=6)	9/10 found the e-consult function of added value Most important reason for not using it was that they had no questions.	Reasons for not having a telephone appointment: Protocol violation (67.3%) Complication (22.4%) Procedure (2.0%)	Mean score of 7.3 (1-10) Most reported reasons for not using it: Connection problems (n=10) No advantages (n=3) Complications (n=3)	**Participants' attitudes**
61%	**60%**	**62%**	**65%**	**61%**	**Implementation score**

Mobile Phone App

Dose Delivered

Of the 173 participants in the intervention group, 159 (91.9%) had a mobile phone or tablet. All these 159 participants received an account and an information brochure containing instructions on how to download and use the app.

Dose Received

Of the 135 participants who received an account for the app and completed the questionnaire 3 months after surgery, 67 (49.6%) answered that they had used the app. A total of 16 participants (23.9%) had used the app only several times, 3 participants (5%) weekly, and 48 participants (72%) several times a week or on a daily basis.

Participants' Attitudes

Participants assessed the app with a mean score of 7.6 on a scale of 1 to 10. Reasons for not using the app were that participants reported that they did not see the added value (n=24), were not able to download the app (n=23), had no mobile phone (n=6), had forgotten it (n=4), suffered from complications (n=4), had no need for it because they had no complaints (n=3), had no time (n=3), or did not find the information they were looking for (n=1). In addition, qualitative data were collected regarding participants' experiences with using the app. Participants who used the app found it a convenient tool. Several aspects of the app were mentioned as being helpful. One participant stated the following:

> *The overview of the convalescenceplan in the app was very useful and gave a good picture about what to expect, I resumed my activities quicker because of the app.* [Female, 46 years old, laparoscopic adnexal surgery]

Another participant stated the following:

> *The recovery monitor in the app which gives feedback on the speed of my recovery in relation to the convalescenceplan provided me with support and made me feel comfortable.* [Female, 38 years old, laparoscopic cholecystectomy]

One participant stated the following:

> *The app was a convenient tool in comparison to the website, because you do not have always your computer quickly available.* [Female, 32 years old, laparoscopic cholecystectomy]

eConsult

Dose Delivered

All participants of the intervention group who received an account for the website were automatically provided with the possibility to ask questions to their health care provider by an eConsult (n=172).

Dose Received

A total of 12 participants (7.0%) made use of the eConsult function.

Fidelity

All 12 questions were answered by the health care providers. Mean time between asking the question and getting a reply was 37 hours.

Participants' Attitudes

The participants who were interviewed mentioned that they had found the eConsult function of added value. The reasons given were that it was an easy or a quick way to ask questions and that they could ask questions while at work. One participant stated the following:

> *The eConsult function was of added value to me, because I have a busy job so I had no time to call the hospital during office hours. Now I could ask my questions after office hours.* [Male, 54 years old, hernia inguinal surgery]

However, the participants did not find it useful for all types of questions. One participant stated the following:

> *The eConsult function on the website is an interesting function, mainly when a quick response could be provided. When you have to wait more than a few days for a response, it will be useless. In addition, for urgent questions (for example high fever) I would have called anyway.* [Female, 38 years old, laparoscopic cholecystectomy]

Another participant stated the following:

> *For more complex questions I would have called the hospital because typing emails is not my strongest point.* [Male, 60 years old, hernia inguinal surgery]

Most of the participants who used the eConsult explicitly mentioned that they would use it again in the future. Most of the eConsult users said that they would not prefer to use the eConsult instead of a (telephone) appointment with the physician, but as an extra facility only. One participant stated the following:

> *In my opinion the eConsult should not replace the appointment in the outpatient clinic. Personal contact with my doctor is important for me. However, a combination of both would be perfect.* [Female, 46 years old, laparoscopic adnexal surgery]

Another participant stated the following:

> *The eConsult should not necessarily replace the appointment in the outpatient clinic. It should be the patient's choice whether or not he or she prefers to have an appointment.* [Female, 40 years old, laparoscopic cholecystectomy]

All the participants who had not used the eConsult function mentioned that they had not used it because they had no questions or complaints. Most of them mentioned that they would have used it if they had questions, but one participant mentioned that she would rather have called the hospital in that case:

> *I have not used the eConsult function because I had no questions or complaints. However, if I had had any questions I rather would have called the hospital*

> *because my question would have been quicker answered.* [Female, 44 years old, laparoscopic cholecystectomy]

Telephone Appointment

Dose Delivered

A total of 89 participants of the intervention group (51.4%) were offered a telephone appointment at the moment of discharge from the hospital. In 25.5% of the participants, it was unclear whether or not the telephone appointment was provided. The remaining 23.1% (n=40) participants were not offered a telephone appointment.

Dose Received

A total of 87 participants (97.8%) received their telephone appointment. The reasons for not receiving the appointment were that that the patient had complaints; therefore, the telephone appointment was replaced by a visit to the outpatient clinic (n=1), and 1 patient requested for an appointment in the outpatient clinic instead of a telephone appointment. The other 61 participants of the intervention group had no postoperative appointment at all (n=37) or had an appointment in the outpatient clinic only (n=49).

Fidelity

A total of 24 participants visited the outpatient clinic in addition to their telephone appointment. In 10 participants, this was decided during the telephone appointment, 8 participants visited the outpatient clinic before the telephone appointment because of a complication or complaints, and the reason was unclear in 6 participants.

Participants' Attitudes

There were 49 participants who only came back at the outpatient clinic and thus did not receive a telephone appointment. In 67% of the participants, this was because of a protocol violation, in 22% because of a complication or complaints, in 6% because a procedure had to be performed, and in 2.0% because of a fertility appointment.

Activity Tracker

Dose Delivered

Of the 173 participants from the intervention group, 124 (71.2%) received an activity tracker. Of the other 55 participants, 14 had no mobile phone and 35 had a mobile phone that was not compatible with the activity tracker.

Dose Received

Of the 124 participants who received an activity tracker, 86 (69.4%) connected the activity tracker to their mobile phone.

Fidelity

A total of 63.7% of the participants who received an activity tracker have used the activity tracker.

Participants' Attitudes

The activity tracker was assessed with a mean score of 7.3 on a scale of 1 to 10. Reasons for not using the activity tracker were as follows: problems with connecting the activity tracker to their phone (n=10), did not see the added value (n=3), suffered from complications (n=3), had no need for it because they had no complaints (n=1), or because the patient felt too sick to use the activity tracker (n=1). Four out of the 12 participants who were interviewed had not used the activity tracker. Reasons were that their mobile phone was not compatible with the activity tracker or that they felt no need to do it. One participant stated the following:

> *I felt no need to connect the activity tracker to my phone, that was too much hassle.* [Male, 60 years old, laparoscopic hernia inguinal surgery]

Most of the participants who used the tracker found it to be a convenient and interesting tool. One participant stated the following:

> *The activity tracker was a motivator to be more active. It was useful to monitor my movements. However I had to remind myself to wear the activity tracker daily.* [Female, 48 years old, laparoscopic adnexal surgery]

Another participant stated the following:

> *The activity tracker was a nice additional tool. It is nice to track how active you are on a day. Sometimes my recovery (displayed on my activity tracker) turned out to be faster than what I was thinking. When I saw for example my activity status of two weeks earlier compared to my current status, I realized that my recovery was going faster than I expected. It was a motivational tool for me.* [Female, 40 years old, laparoscopic cholecystectomy]

One participant stated the following:

> *The activity tracker was easy in use and it was very nice to track my activities. It worked motivating for me. I have moved more to reach my goal. [Female, 32 years old, laparoscopic cholecystectomy].*

Implementation Scores

The implementation scores of the different functions of the intervention are presented in Figure 3. They coincide very closely, ranging between 60% and 65%.

Discussion

Principal Findings

In this process evaluation, the implementation process of a perioperative eHealth intervention, comprising a website, app, activity tracker, eConsult function, and a telephone appointment 2 weeks after surgery, was evaluated. The implementation scores of the different functions of the intervention were fair and ranged between 60% and 65%. The website, app, and activity tracker were assessed with a mean score of 7.3 to 7.6 on a scale of 1 to 10. Twelve study participants were interviewed about the eConsult function; almost all rated it as being of additional value when combined with the usual care.

Interpretation of the Results

The implementation scores were fair, which was caused by the fact that some of the functions of the intervention scored surprisingly low regarding the components of the Linann and Steckler model. In our opinion, there may be three possible reasons for this. First, there was a lack of continuity in providing the intervention to the participants. Because the 173 participants from the intervention group were included in 7 different centers, a mean of 25 participants were included per center in a 1 year period. It is likely that this low volume of patients would have caused the fair implementation and that if the intervention would be implemented in clinical practice outside study setting and, as a consequence, the intervention would be provided to every patient routinely, scores will be much higher. In our opinion, this lack of routine will be the major explanation for the fact that only 25.2% of the convalescence plans were electronically approved by the specialist and only 51.4% were offered a telephone appointment. Another possible explanation for the fair implementation scores is that there is really no need for this specific function. We think that this could be the case in the eConsult function, which was only used in 7% of the participants to whom it had been offered. In the additional interviews we performed, almost all participants answered that they would have used the function in the case they had questions; however, they had no questions. A final explanation could be that participants were hampered by technical barriers to use the intervention. Especially for the activity tracker, the interviews showed that for some participants the different steps that had to be undertaken to install the tracker were a barrier for using it. This could be overcome by a helpdesk providing assistance in this. However, the procedure itself can also be simplified and more easy to use. On the basis of this finding, we can improve the procedures related to installing the activity tracker.

Comparison With Prior Work

In 2014, Bouwsma et al published a process evaluation about the eHealth intervention, which was the base for the development of the eHealth intervention evaluated in this study [15]. Bouwsma reported an implementation score of the eHealth intervention of 80.3%, which was between 15% and 20% higher than the implementation scores of our study. There are some possible explanations for the difference in scores. First of all, and in our opinion the most important one, is the degree of involvement of the researcher in both studies regarding motivating the study participants to use the intervention. For example, the eHealth intervention which was evaluated in the process evaluation of Bouwsma et al was provided to the study population, and when the research team signaled that is was not being used by the study participant, the researcher contacted the study participant to offer assistance. In addition, when the convalescence plan was not approved by the medical specialist, the specialist was contacted by the researcher to bring it to his or her attention. In this study, we tried to limit the involvement of the researcher to a minimum to have a realistic perspective on the actual implementation, including the potential barriers. After the intervention was delivered to the participant, the research team only provided assistance when the study participant contacted them. It was decided to do so as we wanted to create a situation that was most comparable with the situation in which the intervention would be implemented in the future with a helpdesk (outside study setting). In our opinion, this is of great importance as the implementation of eHealth interventions has proven to be a difficult process; so, when we aim to evaluate the barriers of implementation, we should evaluate the intervention in a situation that is as similar to the future situation as possible.

Strengths and Limitations

One of the strengths of the study was the extensiveness in which the process evaluation was performed. This is because the individual functions of the intervention were evaluated separately. By evaluating the individual functions of the intervention, important information was generated according as to what makes the intervention more or less effective, which can be of assistance in the future for the purpose of adapting the intervention. Another strength of the study is the high response rate (84.4%) to the questionnaire that was used to assess participants' attitudes regarding the intervention 3 months after surgery. In addition, the data collection process consisted of several components. Quantitative data containing objective data from a weblog and a logistic database as well as more subjective data assessed by questionnaires were collected, and in addition, qualitative data were collected by interviews with a sample of the study population. However, qualitative data were only collected in a small subsample (n=12) of the study population, and therefore, these data were only presented as an example descriptively and should be interpreted with caution. This study also has some limitations. The most important one is the fact that we have not collected information regarding reasons for nonparticipation. As 60% declined to participate, it would be very valuable to know whether this was because of the study setting and associated burden or because of the fact that the patients had no need for a perioperative eHealth intervention. In our opinion, it is not likely that the latter reason was the major reason for nonparticipation. We performed a survey study 1 year earlier in one of the hospitals that also participated in this study. In this study, patients who had undergone adnexal, hernia inguinal, or a cholecystectomy were also included, and 78% of them indicated that they had felt the need for an eHealth program during their perioperative course [16]. Second, although we think that it is a strength that we evaluated the different functions of the intervention separately, it was difficult to define some components of the Linann and Steckler model regarding the functions of the intervention, for example, the definition of "dose received" of the eConsult function. The nominator was defined as *Participants who asked a question on the Web portal* and the denominator as *Participants who received an account for the Web portal.* Ideally, the denominator would be only the participants who had a question; however, we did not measure this. A final limitation of the study is the questionable manner in which the implementation scores are calculated. This is well illustrated by the fact that we have calculated nearly five identical implementation scores, whereas the individual components of each function of the intervention differed considerably. For example, dose received was 7.0% for the eConsult function and 97.8% for the telephone appointment; however, the implementation scores were nearly the same (62% and 65%, respectively). We used the averaging approach; however, Baranowski et al recommend that the implementation score has to be the result of the product of reach, dose, and fidelity [17].

As the calculation of the implementation score is doubtful, we should be careful with interpreting these scores.

Clinical Implications and Future Research

This study has several important implications. First, the results are of great relevance when interpreting the results regarding the effectiveness evaluation of this study, which will become available in the future. Second, it may be helpful for future research regarding the implementation of these types of eHealth interventions. Unless some eHealth interventions are proven to be effective, the usage in daily practice of the intervention fails most of the time. Therefore, research evaluating the barriers and facilitators for implementing eHealth interventions should be carried out. In this study, we evaluated the feasibility of the intervention. Another study is needed to evaluate implementation barriers in daily practice on a wide scale, because in this study, the intervention was applied in a study setting. This may have influenced the results, mainly the results regarding the dose-delivered component, because the intervention was delivered by the researcher who was likely more involved in the process than, for example, a health care provider who had to deliver the intervention in the future outside the research setting. Finally, the study setting may also have influenced the results, as only 39.9% of the assessed participants gave consent to participate and likely the participants who participated were not a good reflection of the overall population. More qualitative research should therefore be performed in the future focusing on all stakeholders, such as patients, caregivers, and policymakers, that are not involved in an effectiveness study.

Conclusions

In conclusion, participants were overall satisfied with the intervention. However, the implementation scores of the different functions of the intervention were fair. More research is needed to evaluate the barriers and facilitators for implementation of this perioperative eHealth intervention before it can be implemented outside the study setting.

Acknowledgments

This study was carried out with funding from ZonMw, an organization for health research and development in the Netherlands (project number 837002409). The authors would like to thank all departments of Surgery and Obstetrics and Gynecology in the participating hospitals: Diakonessenhuis Utrecht, Flevo Ziekenhuis, Maxima Medisch Centrum, Meander Medisch Centrum, Onze Lieve Vrouwe Gasthuis lokatie Oost, and Spaarne Gasthuis.

Conflicts of Interest

AC, HS, PS, and PD have no conflicts of interest. EM, JH, and JA are the developers of the care program under study. JA and JH intend to set up a spin-off company concerning the implementation of a mobile app and the IKHERSTEL intervention in the Netherlands. JH received grants from NWO, ZonMw, and Samsung during the study and received a fee from Olympus outside the submitted work. HB received personal fees from Olympus, Stryker, and Medtronic and from Applied Medical outside the submitted work. JA holds a chair in Insurance Medicine paid by the Dutch Social Security Agency, he is stockholder of Evalua, and he received grants from ZonMw/NWO, Instituut Gak, UWV, SZW, VWS, Pfizer, Achmea, and CVZ/Zorg Instituut outside the submitted work.

Multimedia Appendix 1

Description of the outcome measures.

[PDF File (Adobe PDF File), 31KB - jmir_v20i1e1_app1.pdf]

Multimedia Appendix 2

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 408KB - jmir_v20i1e1_app2.pdf]

References

1. American Hospital Association. Trends Affecting Hospitals and Health Systems: Utilization and volume URL: http://www.aha.org/research/reports/tw/chartbook/index.shtml.[accessed [accessed 2017-11-30] [WebCite Cache ID 6vMGEvZnE]
2. Brölmann HA, Vonk Noordegraaf A, Bruinvels DJ, de Vet RH, Dirksz AA, Huirne JA. Can prolonged sick leave after gynecologic surgery be predicted? An observational study in The Netherlands. Surg Endosc 2009 Oct;23(10):2237-2241 [FREE Full text] [doi: 10.1007/s00464-008-0287-0] [Medline: 19118421]
3. Clayton M, Verow P. Advice given to patients about return to work and driving following surgery. Occup Med (Lond) 2007 Oct;57(7):488-491. [doi: 10.1093/occmed/kqm063] [Medline: 17906266]
4. Johansen P, Al-Khafagi SK, Thøstesen LM, Lauszus FF, Rasmussen KL. [Analysis of need for sick leave after hysterectomy]. Ugeskr Laeger 2008 Apr 21;170(17):1465-1468. [Medline: 18462627]

5. Tran TT, Kaneva P, Mayo NE, Fried GM, Feldman LS. Short-stay surgery: what really happens after discharge? Surgery 2014 Jul;156(1):20-27. [doi: 10.1016/j.surg.2014.03.024] [Medline: 24856316]
6. Vonk Noordegraaf A, Anema JR, Louwerse MD, Heymans MW, van Mechelen W, Brölmann HA, et al. Prediction of time to return to work after gynaecological surgery: a prospective cohort study in the Netherlands. BJOG 2014;121(4):487-497. [doi: 10.1111/1471-0528] [Medline: 24245993]
7. Bouwsma EV, Anema JR, Vonk Noordegraaf A, Knol DL, Bosmans JE, Schraffordt KSE, et al. The cost effectiveness of a tailored, web-based care program to enhance postoperative recovery in gynecologic patients in comparison with usual care: protocol of a stepped wedge cluster randomized controlled trial. JMIR Res Protoc 2014 Jun 18;3(2):e30 [FREE Full text] [doi: 10.2196/resprot.3236] [Medline: 24943277]
8. Vonk Noordegraaf A, Huirne JA, Brölmann HA, Emanuel MH, van Kesteren PJ, Kleiverda G, et al. Effectiveness of a multidisciplinary care program on recovery and return to work of patients after gynaecological surgery; design of a randomized controlled trial. BMC Health Serv Res 2012 Feb 01;12:29 [FREE Full text] [doi: 10.1186/1472-6963-12-29] [Medline: 22296950]
9. Vonk Noordegraaf A, Anema JR, van Mechelen W, Knol DL, van Baal WM, van Kesteren PJ, et al. A personalised eHealth programme reduces the duration until return to work after gynaecological surgery: results of a multicentre randomised trial. BJOG 2014 Aug;121(9):1127-35; discussion 1136 [FREE Full text] [doi: 10.1111/1471-0528.12661] [Medline: 24511914]
10. van der Meij E, Huirne JA, Bouwsma EV, van Dongen JM, Terwee CB, van de Ven PM, et al. Substitution of usual perioperative care by eHealth to enhance postoperative recovery in patients undergoing general surgical or gynecological procedures: study protocol of a randomized controlled trial. JMIR Res Protoc 2016 Dec 21;5(4):e245 [FREE Full text] [doi: 10.2196/resprot.6580] [Medline: 28003177]
11. Ross J, Stevenson F, Lau R, Murray E. Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). Implement Sci 2016 Oct 26;11(1):146 [FREE Full text] [doi: 10.1186/s13012-016-0510-7] [Medline: 27782832]
12. Murray E, Burns J, May C, Finch T, O'Donnell C, Wallace P, et al. Why is it difficult to implement e-health initiatives? A qualitative study. Implement Sci 2011;6:6 [FREE Full text] [doi: 10.1186/1748-5908-6-6] [Medline: 21244714]
13. Eysenbach G. 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]
14. Linnan L, Steckler A. A process evaluation for public health interventionsresearch± an overview. In: Proces evaluation for public health interventions and research. 1st edition. San Francisco, CA: Jossey-Bass Incorporated Publishers; 2002:1-23.
15. Bouwsma E, Vonk Noordegraaf A, Szlávik Z, Brölmann H, Emanuel M, Lips J, et al. Process evaluation of a multidisciplinary care program for patients undergoing gynaecological surgery. J Occup Rehabil 2013 Sep 22;118(13):1557-1567 [FREE Full text] [doi: 10.1007/s10926-013-9475-4] [Medline: 21895950]
16. van der Meij E, Bouwsma EV, van den Heuvel B, Bonjer HJ, Anema JR, Huirne JA. Using e-health in perioperative care: a survey study investigating shortcomings in current perioperative care and possible future solutions. BMC Surg 2017 May 23;17(1):61 [FREE Full text] [doi: 10.1186/s12893-017-0254-6] [Medline: 28535763]
17. Baranowski T, Stables G. Process evaluations of the 5-a-day projects. Health Educ Behav 2000 Apr;27(2):157-166. [doi: 10.1177/109019810002700202] [Medline: 10768797]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
eConsult: electronic consultation
eHealth: electronic health
SD: standard deviation

Edited by G Eysenbach; submitted 04.07.17; peer-reviewed by W Leclercq, R Staerkle; comments to author 27.07.17; revised version received 17.08.17; accepted 15.09.17; published 02.01.18.

Please cite as:
van der Meij E, Huirne JAF, ten Cate AD, Stockmann HBAC, Scholten PC, Davids PHP, Bonjer HJ, Anema JR
A Perioperative eHealth Program to Enhance Postoperative Recovery After Abdominal Surgery: Process Evaluation of a Randomized Controlled Trial
J Med Internet Res 2018;20(1):e1
URL: http://www.jmir.org/2018/1/e1/
doi:10.2196/jmir.8338
PMID:29295808

Original Paper

Turning Good Intentions Into Actions by Using the Health Action Process Approach to Predict Adherence to Internet-Based Depression Prevention: Secondary Analysis of a Randomized Controlled Trial

Anna-Carlotta Zarski[1,2], MSc; Matthias Berking[1], PhD; Dorota Reis[3], PhD; Dirk Lehr[2], PhD; Claudia Buntrock[1], MSc; Ralf Schwarzer[4], PhD; David Daniel Ebert[1], PhD

[1]Friedrich-Alexander-University Erlangen-Nürnberg, Erlangen, Germany
[2]Leuphana University Lüneburg, Lüneburg, Germany
[3]University Koblenz-Landau, Landau, Germany
[4]SWPS University of Social Sciences and Humanities, Warszawa, Poland

Corresponding Author:
Anna-Carlotta Zarski, MSc
Friedrich-Alexander-University Erlangen-Nürnberg
Nägelsbachstraße 25a
Erlangen,
Germany
Phone: 49 9131 85 67570
Email: Anna-Carlotta.Zarski@fau.de

Abstract

Background: Many individuals engaging in Internet-based interventions fail to complete these treatments as intended. The processes responsible for treatment adherence in Internet-based interventions are still poorly understood.

Objective: The aim of this study was to investigate to what extent adherence in an Internet-based intervention can be predicted by motivational and volitional factors outlined in the health action process approach (HAPA).

Methods: This study investigated motivational and volitional factors included in HAPA in a randomized controlled trial to predict treatment adherence of N=101 individuals with subclinical depression in the intervention group of a depression prevention intervention (GET.ON Mood Enhancer). Adherence was operationalized as the number of completed treatment modules. Using longitudinal structural equation modeling, HAPA variables (motivational, maintenance, and recovery self-efficacy, outcome expectancies, intention, and planning) were assessed at baseline and their associations with adherence 7 weeks later.

Results: Planning predicted adherence. Better planning was, in turn, associated with higher levels of maintenance self-efficacy, and the latter significantly affected treatment adherence via planning. The other hypothesized direct associations were not significant. In total, the HAPA variables accounted for 14% of variance in treatment adherence.

Conclusions: Planning emerged as the strongest predictor of treatment adherence in highly motivated participants in an Internet-based intervention out of all HAPA variables investigated. Findings are in line with the hypothesis that planning facilitates the translation of good intentions into actions. The findings imply that systematically fostering planning skills and maintenance self-efficacy prior to or during Internet-based interventions would help participants to successfully complete these treatments.

Trial Registration: German Clinical Trials Register DRKS00005973; https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00005973 (Archived by WebCite at http://www.webcitation.org/6uxCy64sy).

(J Med Internet Res 2018;20(1):e9) doi:10.2196/jmir.8814

KEYWORDS

health action process approach; adherence; Internet intervention; depression prevention

Introduction

Internet-based interventions have been shown to prevent the incidence of mental health disorders such as depression or anxiety and reduce associated symptom severity [1,2]. Internet-based interventions used to prevent mental health disorders have several advantages. First, they are readily accessible at any time and place. Second, individuals can choose to remain anonymous and thus avoid stigmatization. Third, individuals tend to have more active roles in (guided) self-help interventions, and for this reason, it might be easier for them to integrate the newly acquired skills in their day-to-day lives. Fourth, individuals can work at their own pace and go through materials as often as they want [3].

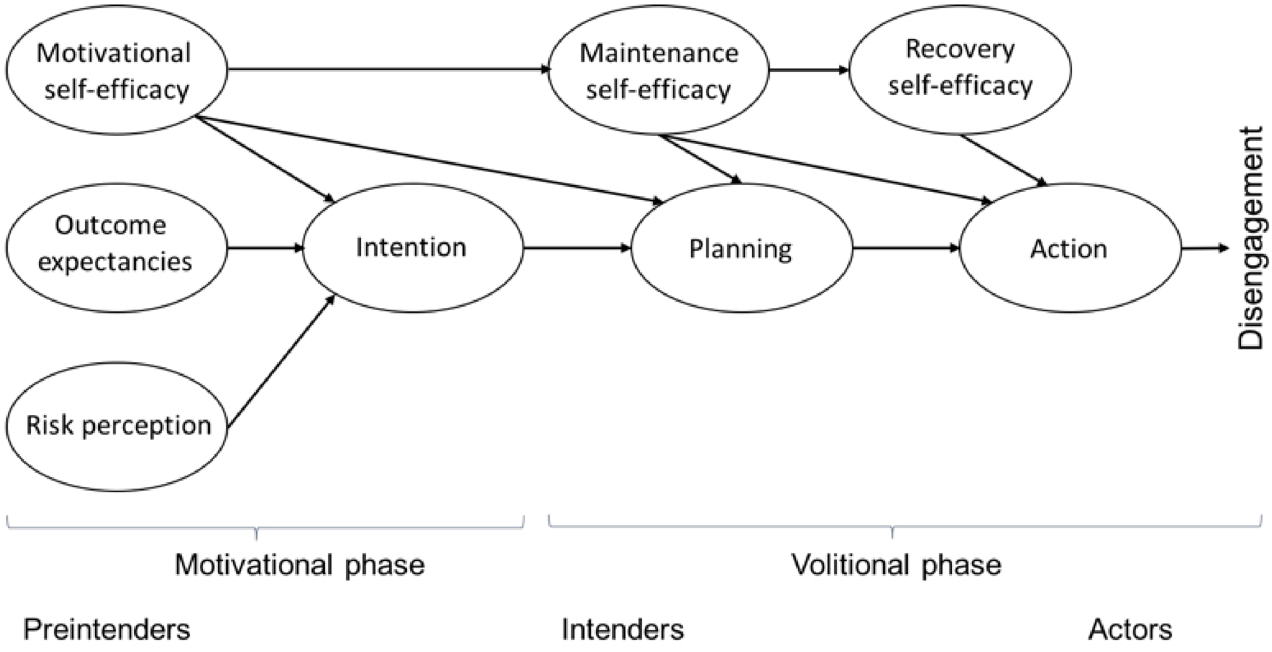
The high autonomy and flexibility of Internet-based interventions facilitate low-threshold access to treatment but also place high self-regulatory demands on participants and thereby entice treatment cessation [4]. Low treatment adherence rates can, in turn, reduce the effectiveness of Internet-based interventions substantially [5,6]. Many individuals struggle to begin or complete Internet-based interventions despite being highly motivated. This suggests that good intentions do not guarantee sustained adherence or address how to adequately deal with barriers associated with Internet-based interventions such as allocating time to work on the training modules [7]. It also shows the need to clarify which factors and processes determine whether participants can put their intentions into action in order to complete Internet-based interventions.

Concerning sociodemographic characteristics and symptom severity and their influence on treatment adherence, few clear predictors have emerged. Female gender, for example, has been shown to be a predictor of higher treatment adherence in contrast to, for example, level of education, marital status, employment, or ethnicity, which have not been found to be associated with adherence [8]. For age and baseline symptom severity, the results were inconsistent [8]. The findings of previous studies can, however, explain the variance in adherence only to some extent [9-12], with findings of, inter alia, 11% [13]. There is empirical evidence showing that the intrinsic motivation to commence an intervention, treatment expectancy, and the ability to focus on future goals influence adherence and attrition in Internet-based interventions [7,14-16]. However, only a few studies used a theoretical framework to explain adherence in Internet-based intervention [14].

As suggested by the theoretical framework of the health action process approach (HAPA), the gap between intention and behavior can be explained by volitional factors [17,18]. Describing 2 phases and 3 stages, HAPA can explain why individuals adopt and maintain a wide range of health behaviors [17,19]. As suggested by this model, a motivational phase, in which the intention to adopt a certain kind of health behavior is developed, is followed by a volitional phase, in which behavior is planned, prepared, and executed (see Figure 1) [20]. HAPA also assumes that individuals pass through different stages such as preintention, intention, and action when adopting new behavior [21,22]. In the motivational phase, individuals are seen as preintenders if they have not yet decided to adopt a new behavior. In the volitional phase, individuals are classified as either intenders who have decided to adopt the target behavior or as actors already performing the behavior [21,23].

The influence of variables of the motivational and volitional phases on health behavior depends on the stage individuals are currently in (see Figure 1) [24,25]. In the motivational phase, to develop an intention is assumed to be influenced by outcome expectancies, risk perceptions, and motivational self-efficacy [26-29]. Outcome expectancies are a person's positive and negative expectations concerning the consequences of adopting a behavior, and they have been shown to be the strongest predictor in the motivational phase [20].

Figure 1. The health action process approach model.



Risk perception is a measure of perceived vulnerability in terms of health impairment, and compared to outcome expectancies, this has shown to be a weaker predictor of intention [20,26,30]. Motivational self-efficacy, which can be defined as the belief in one's ability to perform the targeted behavior, is regarded as the second best predictor of behavioral intentions [27]. By expressing an explicit behavioral intention, individuals are motivated to act, although such motivation does not necessarily need to be translated into actual behaviors, given the likelihood that barriers emerge rendering the intention instable. Thus, an intention may be seen as a distal antecedent of action. It is assumed that initial motivation at the onset of an intervention makes a difference for all subsequent processes, thus, influencing the likelihood of planning as well as the eventual success of the intervention as reflected by higher adherence levels.

In the volitional phase, individuals initiate and maintain the target behavior [31]. In this phase, maintenance self-efficacy and recovery self-efficacy have shown to be crucial for individuals facing imminent barriers and for those coming to terms with relapses, respectively [32,33]. Moreover, planning has been shown, in line with HAPA, to be a mediator between intentions and behaviors and to further facilitate the translation of intentions into actions [30,34-37]. Planning is regarded as a prospective self-regulatory skill where an individual specifies the situational context in which one will enact to ensure that behavioral performance is achieved. Planning requires a mental representation of how to achieve some future outcome that allows the individual to link the intended behavior with a particular context for its enactment, thus connecting the individual with good opportunities to act. Planning may also include the anticipation of barriers and the generation of alternative behaviors to overcome those [38].

Currently, it is unclear to what extent motivational and volitional variables included in HAPA can also explain the intention-action gap in the field of adherence in Internet-based interventions. To shed light on these factors and processes in Internet-based interventions, this study explores the intention-behavior gap by assessing motivational and volitional adherence predictors based on the HAPA model for individuals showing subclinical symptoms of depression but not fulfilling the criteria for a major depressive disorder. Treatment adherence is operationalized here by the number of completed treatment modules. First, according to HAPA, it can be expected that differences in treatment adherence between the participants are mainly due to volitional factors such as planning because participants can already be classified as intenders due to their decision to take part in an Internet-based intervention for depression prevention [21]. Accordingly, motivational self-efficacy and outcome expectancies should not predict intention. Second, it was hypothesized that higher levels of planning, maintenance, and recovery self-efficacy in the volitional phase should explain higher rates of treatment adherence rates in this sample. Hypotheses were tested using a structural equation model based on a longitudinal research design.

Methods

The study outlined below was described in greater detail elsewhere [39]. Data for the secondary analyses were collected in a randomized controlled trial (RCT) evaluating a guided Internet-based intervention for depression prevention (GET.ON Mood Enhancer) comparing an intervention group to a waitlist control group. Participants assigned to the waitlist control group did not have access to the intervention during the first 3 months after randomization. For the analyses in this study, assessment took place at baseline (T1) and at posttreatment 7 weeks after randomization (T2). The study was approved by the medical ethics committee of the Leuphana University of Lüneburg (reference number: Ebert201404_Depr) and registered with the German Clinical Trials Register [DRKS00005973].

Sample

The analyses were conducted using the intervention group sample (n=102), which was given access to the Internet-based intervention directly after randomization. One participant of the intervention group was excluded because of missing data in the HAPA questionnaire at baseline, resulting in a sample of n=101 for this analysis. Applicants were included in the study if they (1) had a subthreshold depression (Center for Epidemiological Studies Depression Scale [CES-D] ≥16), (2) were 18 years or older, (3) had Internet access, (4) were willing to give informed consent, and (5) did not show a notable suicidal risk (Beck Depression Inventory item 9 >1) [40-42]. Exclusion criteria were (1) a current major depressive episode as defined by the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision* criteria assessed with the Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition) Axis I Disorders [43], (2) a major depressive episode in the past 6 months, (3) bipolar disorder, (4) psychotic disorder, (5) currently receiving psychotherapy or having received psychotherapy for any kind of mental health disorder in the past 6 months, or (6) being on a waiting list for psychotherapy for any kind of mental health problems. All applicants could use routine care (eg, they could visit their general practitioners).

Procedure

Potential participants were recruited from the general population with the help of a large German health insurance company (Barmer Gmünder Ersatzkasse), and the study was also announced in newspaper articles, on-air media, and related websites. Individuals interested in participating in the study applied online on the website designed for this study (www.geton-training.de) by submitting their email address or by sending an email to the research team. Applicants then received an information letter via email with detailed information on the intervention and the study. In this letter, they were informed that they could withdraw from the intervention or study at any time without any negative consequences. Applicants were then screened for study inclusion and exclusion criteria, and those who met all of the criteria considered during this screening process were scheduled for the semistructured clinical interview conducted by telephone [44,45]. Individuals who met none of the exclusion criteria after the interviews, completed the baseline assessment, and returned the informed consent form via mail or email entered the study. They were then randomly allocated to either the intervention group or the waitlist control group. Balanced block randomization took place at an individual level in blocks of 12 to maintain a ratio of 1:1 between the 2 study groups. This process was completed using an automated computer-based random numbers table and completed by a researcher not involved in the study.

Intervention

The intervention was described in greater detail by Buntrock and her colleagues [1,39,46]. It is based on behavior therapy and problem-solving therapy and consists of 6 modules and 1 booster session. Additional elective modules integrated in the last 3 sessions are directed at sleep hygiene, relaxation techniques, and dealing with worrying thoughts. Each session can be completed in approximately 30 to 60 minutes. Participants were advised to attend a maximum of 2 sessions per week but at least 1. Consequently, the training takes about 3 to 6 weeks plus a booster session 4 weeks after the end of the training. Lessons consist of general text-based information, testimonials, interactive elements such as exercises, and other content such as mp3 audio files, video clips, and downloadable work sheets. The intervention was conceptualized as guided self-help, and intervention elements such as self-gratification were included to support participants during the self-help process and encourage them to continue treatment. The training was gradually adjusted to the specific needs of individual participants based on their responses to and choices of different options. Participants were encouraged to keep a daily online training diary to monitor their mood and reflect on mood-enhancing activities. One key feature of the intervention is the focus on homework assignments, which allowed participants to integrate newly acquired coping skills and techniques into their daily lives. A secure Web-based platform (Advanced Encryption Standard [AES] 256-bit encryption) was used for the training. Participants accessed the intervention on the platform using their email addresses and passwords that they had created. If desired, participants received a set of about 42 standardized automatic motivational text messages including descriptions of short exercises on their mobile phones.

Adherence-Focused Guidance

To support their adherence to the training, participants received guidance by an electronic coach (eCoach) using an adherence-focused guidance concept, described in detail elsewhere [12,47]. Adherence-focused guidance consisted of adherence monitoring and feedback on demand. Adherence monitoring included checking module completion on a regular basis and sending reminders in case participants had not completed at least 1 module within 7 days. Feedback on demand included giving participants the opportunity to contact the eCoach and receive individual support or feedback on training modules within 48 hours. Only a few participants (6/101, 5.9%) requested feedback, resulting in 15 instances of content feedback for the entire sample. This corresponds to an average of 0.15 feedback demands per participant (range 0-5, standard deviation [SD] 0.71). Checking module completion and providing reminders, then, accounted for most of the time spent per participant. The eCoaches were trained psychologists who followed feedback guidelines defined in the standardized manual for the intervention, in accordance with the supportive accountability model [48]. The supportive accountability model assumes that human support in the context of Internet-based interventions increases adherence rates because participants tend to develop a sense of commitment toward an eCoach, who is perceived as trustworthy, benevolent, and knowledgeable.

Measures

Self-report measures for the present analyses were collected at baseline and at posttreatment also using the secured online-based assessment system (AES 256-bit encrypted).

Sociodemographic Information and Depression

Data on sociodemographic information and depression were collected at baseline. Depressive symptom severity was measured by the self-report CES-D, the clinician-rated Hamilton Rating Scale for Depression (HRSD24), and the Quick Inventory of Depressive Symptomatology–Clinician Rating (QIDS-CR16). A cutoff on the CES-D of 23 is regarded as an indicator of clinically relevant depressive symptoms in German samples [42,49]. The cutoff points of 10, 19, 27, and 35 of the HRSD24 indicate mild, moderate, severe, and very severe depression, respectively [50-52]. The QIDS cutoff points of 6, 11, 16, and 21 represent the thresholds for mild, moderate, severe, and very severe depression, respectively [50].

Health Action Process Approach Measures

The HAPA questionnaire, designed in accordance with the guidelines prepared by Schwarzer [19], was completed at baseline. All items were measured on a 4-point Likert scale ranging from 1=not true at all to 4=exactly true.

Motivational self-efficacy regarding the capability of the participants to complete the training modules including the exercises was measured with 2 items. One item was “I am confident that I am able to complete all 6 modules of the online training and the booster session 4 weeks after completion of the training even if there might be problems.” The Spearman-Brown reliability estimate was reported for the 2-item scale and showed excellent internal consistency (r_s=.91) in this study.

Outcome expectancies were assessed with 2 items measuring the subjective beliefs concerning the positive impact of training adherence on mental health outcomes. One of the items used here was “If I complete 1 module of the online training per week, I will become more resilient in my everyday life.” The Spearman-Brown reliability estimate for this scale showed adequate internal consistency (r_s=.66).

Intention was assessed with 1 item asking participants to what extent they intend to complete all 6 modules of the online training and the booster session 4 weeks after completing the training.

Planning was measured with 4 items, and this variable was used to assess whether participants have made concrete plans when and how they will complete the training, also in case of potential difficulties. One of the items was “I have already made detailed plans how often I will work on the modules during the week.”

In this study, Cronbach alpha was found to be excellent (alpha=.92).

Maintenance self-efficacy was measured with 4 items focusing on potential obstacles during later stages of the online training. Barriers considered in here were no immediate positive effects, technical problems, perceived difficulties, and lack of motivation. "I am confident that I am able to complete 1 module of the online training per week even if I do not see positive effects immediately," was one of the items used in this study. Cronbach alpha was acceptable (alpha=.67).

Recovery self-efficacy was measured with 3 items. The term is used to describe participants' belief that they can deal with failure and continue to work on the training modules after being nonadherent (ie, after postponing concrete plans and not using the intervention for more than a week). One of the items included here was "I am confident that I can continue working on the training modules even if I postpone my detailed plans several times." Cronbach alpha was found to be excellent (alpha=.91) in this study.

Risk perception was not included in the model because this factor is likely to be of minor importance in this study sample. Individuals participating in this prevention intervention are assumed to already perceive a high risk for developing a depression which led to help seeking in an Internet-based intervention for depression prevention.

Adherence Measure

The number of completed treatment modules in the Internet-based intervention for depression prevention, which ranged from 0 to 7 including the 6 core modules and the booster session 4 weeks after treatment completion, was the primary outcome measure in this study and was tracked automatically by the training platform. To complete a module, participants had to respond to all writing tasks and submit the modules to the system. A module completion score of 0 meant that the participant either did not start the intervention or did not finish the first module.

Data Analysis

Reliability and descriptive analyses were performed with SPSS 23 (IBM Corp). Structural equation modeling was applied to assess the HAPA model fit and test the hypothesized associations between the model constructs using the Lavaan package in R (The R Foundation). Maximum likelihood parameter estimation was used with robust (Huber-White) standard errors and a scaled test statistic that is (asymptotically) equal to the Yuan-Bentler test statistic [53]. The structural equation model included the latent exogenous variables outcome expectancies, motivational self-efficacy, maintenance self-efficacy, and recovery self-efficacy; the mediating latent variable planning; and the manifest endogenous variables intention and treatment adherence.

The model fit was assessed with the goodness-of-fit indices chi-square (χ^2), the χ^2 value relative to its degrees of freedom (χ^2/df), the root mean square error approximation (RMSEA), the standardized root mean square residual (SRMR), the comparative fit index (CFI), and the Tucker-Lewis Index (TLI). Adequate model fit was indicated by a nonsignificant χ^2 value, a χ^2/df ratio between 0 and 2, CFI and TLI values greater than .95, RMSEA value below .06, and SRMR values below .08 [54]. Four planning items were combined into 2 parcels and used as indicators for the variable planning [55]. For the variable maintenance self-efficacy, a single higher order factor was specified.

Results

Descriptive Statistics

Participant characteristics at baseline are shown in Table 1; 80.2% (81/101) of the participants were women. Participants had a mean age of 45 years (SD 11.68 ranging from 23 to 75 years) and an above-average level of education (general qualification for university entrance or higher; 83/101, 82.2%). Many participants had prior experience with psychotherapy (43/101, 42.6%), but only a few had taken advantage of health-related trainings (23/101, 22.8%). As indicated by their responses to the CES-D, the participants showed clinically relevant depressive symptoms (mean 26.61, SD 6.51). The results of the clinical interview at baseline showed that participants were, on average, mildly depressed ($\text{mean}_{\text{HRSD}}$ 13.72, SD_{HRSD} 6.24; $\text{mean}_{\text{QIDS}}$ 8.18, SD_{QIDS} 3.63).

As shown in Table 2, participants were, on average, characterized by very high motivational self-efficacy (mean 3.64, SD 0.46) and intention (mean 3.51, SD 0.84) to complete the treatment modules. Participants' expectations concerning the outcome were high (mean 3.21, SD 0.51), and so were the results concerning maintenance self-efficacy (mean 3.36, SD 0.47) and recovery self-efficacy (mean 3.41, SD 0.57). Participants indicated, however, that they had, on average, not made any specific plans when and how they would complete the training (mean 2.44, SD 0.92).

Adherence Rates

Figure 2 depicts the number of completed modules. In total, 5.9% (6/101) of the participants did not start the intervention, whereas 62.4% (63/101) completed all 6 core modules, and 39.6% (40/101) completed all 7 modules including the booster session. On average, participants completed 5.12 modules (SD 2.22, range 0-7).

Table 1. Baseline characteristics of the study population (N=101).

Characteristic	Intervention group (N=101)
Age, years, mean (SD[a])	44.57 (11.68)
Gender, n (%)	
Female	81 (80.2)
Male	20 (19.8)
Ethnicity, n (%)	
White	78 (77.2)
Not reported	23 (22.8)
Relationship, n (%)	
Single	26 (25.7)
Married or cohabited	65 (64.4)
Divorced or separated	9 (8.9)
Widowed	1 (1.0)
Level of education, n (%)	
Low[b]	2 (2.0)
Middle[c]	16 (15.8)
High[d]	83 (82.2)
Employment status, n (%)	
Employed	89 (88.1)
Unemployed or seeking work	2 (2.0)
On sick leave	0
Not employed	10 (9.9)
Gross annual income (Euro), n (%)	
Low (<10,000)	9 (8.9)
Middle (10,000-60,000)	69 (68.3)
High (>60,000)	14 (13.9)
Not reported	9 (8.9)
Experience with health-related trainings, n (%)	
Yes	23 (22.8)
No	78 (77.2)
Experience with face-to-face psychotherapy, n (%)	
Yes	43 (42.6)
No	58 (57.4)
Use of antidepressants, n (%)	7 (6.9)
CES-D[e] sum score, mean (SD)	26.61 (6.51)
HRSD[f] sum score, mean (SD)	13.72 (6.24)
QIDS[g] sum score, mean (SD)	8.18 (3.63)

[a] SD: standard deviation.

[b] bQualifications below a degree from a German secondary school (Realschule).

[c] Degree by a German secondary school or higher; apprenticeship.

[d] General qualification for university entrance or higher.

[e] CES-D: Center for Epidemiologic Studies Depression Scale.

[f] HRSD: Hamilton Rating Scale for Depression.

[g] QIDS: Quick Inventory of Depressive Symptomatology–Clinician Rating.

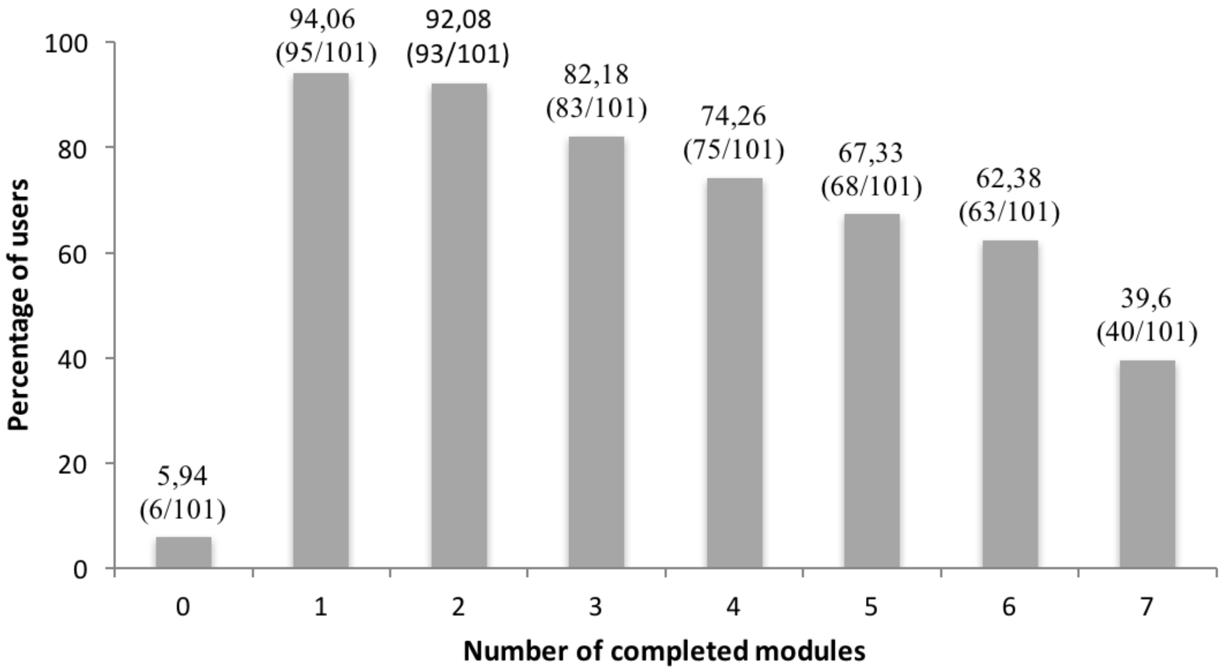
Table 2. Means and standard deviations for the health action process approach variables at baseline.

HAPA[a] variables	Mean (SD[b])	Range
Motivational self-efficacy (2 items)	3.64 (0.46)	3-4
Outcome expectancies (2 items)	3.21 (0.51)	1-4
Intention (1 item)	3.51 (0.84)	1-4
Planning (4 items)	2.44 (0.92)	1-4
Maintenance self-efficacy (4 items)	3.36 (0.47)	2-4
Recovery self-efficacy (3 items)	3.41 (0.57)	1-4

[a] HAPA: health action process approach.

[b] SD: standard deviation.

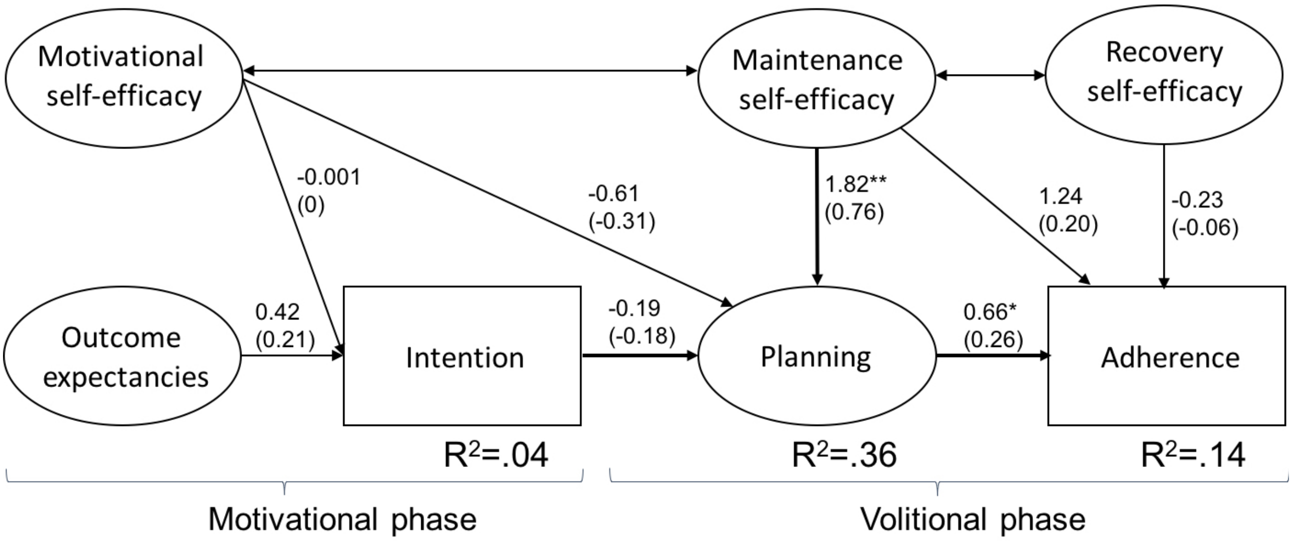
Figure 2. Treatment adherence rates per module.



Health Action Process Approach Model Fit

According to the criteria proposed by Schermelleh-Engel et al [54], the HAPA model for treatment adherence yielded a good fit, χ^2_{82}=106.163, *P*=.038, χ^2/df=1.29, CFI=.96, TLI=.95, RMSEA=.05 (90% CI 0.01-0.08), SRMR=.07, despite the significant results of the chi-square test [54]. Although the chi-square test is often used to evaluate models, this statistic is known to be sensitive to sample size [56]. Figure 3 shows the unstandardized parameter estimates and the standardized parameter estimates (in parentheses). As hypothesized with regard to the volitional phase, higher levels of maintenance self-efficacy predicted more planning (*ß*=0.76, *P*<.001). Planning, in turn, was a significant predictor of treatment adherence at T2 (*ß*=.26, *P*=.04). The other direct associations predicted by HAPA were not significant. Maintenance self-efficacy significantly affected treatment adherence through planning (*b*=1.20, Monte Carlo 95% CI 0.06-2.71). This corresponds to a medium effect of .19 (index of mediation) [57]. The model accounted for 4% of the variance in intention and 36% of the variance in planning. Altogether, 14% of the variance in treatment adherence could be explained.

Figure 3. Prediction of treatment adherence with the health action process approach. Note: upper coefficients are unstandardized parameter estimates; lower coefficients (in parentheses) are standardized parameter estimates. *$P<.05$, **$P<.001$.



Discussion

Principal Findings

Treatment adherence is challenging for many participants using Internet-based interventions: they intend to participate in an intervention, but they either do not start or do not complete it. This study explored this intention-behavior gap by focusing on motivational and volitional factors as explained by HAPA, adapted with regard to treatment adherence.

As hypothesized, volitional processes clearly affected treatment adherence. More specifically, planning emerged as a significant predictor of treatment adherence. This result indicates that planning might point to an underlying mechanism in adherence in Internet-based interventions. This finding is in line with studies showing that interventions that require participants to engage repeatedly in strategic planning may work because mentally linking intentions to specific contexts increases the likelihood of translating intentions into actions [37,58-61]. As also suggested by the results, it might be easier for individuals with high levels of maintenance self-efficacy to engage in planning than for those with low levels, as the former are confident that they can overcome adherence barriers such as technical problems and the absence of immediate feedback [62,63]. Maintenance self-efficacy, in turn, led to higher treatment adherence through planning. It is, however, also possible that, contrary to the relationship suggested by HAPA, planning might affect maintenance self-efficacy. In this study, we evaluated the proposed relationships between the HAPA variables as suggested by the model. To assess the causality and direction of the effect, a longitudinal intervention study controlling for other influential variables or an experimental design would be necessary.

The motivational profile of individuals who decided to use an Internet-based intervention assumed to be intenders was confirmed by the results showing high levels of motivational self-efficacy, outcome expectancies, and intention of participants in this study. They were, however, not predictive of participants' treatment adherence. Thus, in line with HAPA, motivational variables seem to no longer exert influence once participants have developed the motivation and intention to participate in an intervention [64]. In contrast, other studies have found the intrinsic motivation of individuals, their belief in their own ability to complete the intervention, and their expectancies regarding treatment outcome to be associated with adherence in Internet-based interventions [7,14,65-68]. Participants of those studies might have been situated in a preintentional stage due to a less elaborated study inclusion process and less detailed information about the expected commitment prior to intervention start, which might have led to the greater influence of intrinsic motivation and expectations on adherence. This assumption is supported by another study which had an elaborated screening process that also found intentions to use an Internet-based intervention not to be related to actual adherence [7].

This study showed that 14% of the variance in treatment adherence could be explained with motivational and volitional processes. In comparison to other studies evaluating predictors of adherence, the amount of explained variance in this analysis is relevant given the small number of variables significantly influencing treatment adherence in this sample and given the restricted variance in variables due to the homogenous study population [12,69]. In a study evaluating the influence of different guidance formats, gender, age, education, symptom-related factors, and hope for improvement, 9.4% of the variance in treatment nonadherence could be explained [12].

While volitional processes seem to be important mechanisms of treatment adherence to Internet-based interventions, future research needs to investigate to which degree other psychological or social variables additionally influence or moderate treatment adherence. Future studies could consider additional HAPA variables such as action control as well as barriers or facilitators that are closely related to the HAPA constructs (eg, perceived social support, perceived support by an eCoach guiding individuals using an Internet-based intervention, or intervention characteristics such as usability). In general, variables may need to be adjusted to the specific context (ie, adherence to and engagement with Internet-based interventions). Moreover, it would be necessary to test whether

HAPA variables have an incremental influence beyond sociodemographic, disease-related, and intervention variables.

Limitations

This study has several limitations. First, treatment adherence was operationalized as the number of completed treatment modules. Completing a module requires working through different writing tasks, but it is difficult to discern whether participants truly engaged with the content of the intervention or applied what they had learned by completing exercises in their day-to-day lives. Second, there might be a subgroup of participants who discontinued using the intervention because they have already attained their personal treatment goal before the end of the training and did not need to complete the entire intervention [70,71]. Thus, higher treatment adherence might not always be related to better treatment outcome. Third, the variance in treatment adherence was restricted, and the sample was very homogeneous. This may have been due to the elaborate screening process, which required prospective participants to be highly motivated to be considered for this study. Thus, these results may underestimate the effects of the HAPA variables on treatment adherence compared to other studies designed to analyze adherence predictors. It is important to note here that the sample consisted of people with depressive symptoms, who show larger volitional deficits than people with other disorders or healthy ones. The results of this study might also in this respect underestimate the effects of HAPA variables on adherence [62,71]. Fourth, participants also received adherence reminders and feedback on demand; both of these elements have been associated with higher adherence rates [12]. Guidance as a potential influence of adherence was, however, a constant factor among all participants in this study because only a few requested feedback. Therefore, guidance is likely to have affected the level of adherence but not to have led to interindividual differences. Fifth, only one adherence measure was included in the analysis. In future studies, different adherence measures need to be used to collect more data on the quality of engagement with an intervention (eg, number of online training diary entries). Moreover, only self-report measures were used, and the reliability was restricted for some of the constructs. The outcome measure (ie, treatment adherence) was assessed objectively. Sixth, due to feasibility limitations, the HAPA variables were only assessed at baseline and therefore cannot account for individual changes in motivational and volitional attitudes at different stages during the treatment process, although these changes might be relevant for treatment adherence. Future studies should therefore include additional measurement points over the course of the intervention to assess motivational and volitional variables concerning individual sessions. Furthermore, negative outcome expectancies and action control were not included in the HAPA model assessed in this study and should be analyzed with regard to treatment adherence in a next step.

Conclusion and Recommendations for Future Research

As shown in this study, planning was the strongest predictor of treatment adherence and, therefore, should be a key dimension of future Internet-based interventions. Maintenance self-efficacy seems to be a crucial prerequisite in this respect, especially because it allows individuals to overcome potential barriers in the course of treatment. Individuals who had already decided to use the intervention did not seem to need further motivation or positive outcome expectancies at the beginning of this study. Instead, one explanation might be that they might need further support when it comes to detailed planning on how to complete modules on a regular basis to maintain or increase their adherence motivation while using the intervention. Individuals low in planning competences may therefore benefit from identifying possible obstacles and barriers with regard to module completion and develop coping strategies early in the intervention to keep up with module completion. To foster the implementation of these action and coping plans, participants should have the option to formulate if-then plans (eg, “If I do not feel like logging in and completing a module, then I review my treatment goals”). For maintenance self-efficacy purpose, evaluating treatment barriers and developing coping strategies should be repeated throughout the intervention. At the end of each module, participants might also profit from scheduling their next log-in to the intervention for the upcoming week. When individuals do not achieve their personal adherence goals, additional support should be provided to motivate them to retry, choose different coping strategies, or adapt their goals. The value of such tailored strategies in Internet-based preventive interventions to foster volitional competences regarding treatment adherence in individuals should be evaluated systematically in future research. In this respect, it is also important to identify what works best for whom because different features may have different effects on individuals depending on the motivational state they are in.

Two major strengths of this study were its longitudinal design and objective adherence measure. Due to the limitations outlined above, the main findings of this study will, however, have to be confirmed by future research, which will also have to consider other psychological disorders. Using theoretical frameworks such as HAPA when designing interventions and conducting research is important because it allows researchers to test a given theory’s proposed relationships, and, if these can be confirmed, this approach could provide a blueprint for effective future interventions.

Conflicts of Interest

MB, DDE, and DL are stakeholders of the Institute for Online Health Trainings that aims to transfer scientific knowledge related to this research into routine health care.

Multimedia Appendix 1

Correlation matrix and health action process approach item scores.

[PDF File (Adobe PDF File), 70KB - jmir_v20i1e9_app1.pdf]

References

1. Buntrock C, Ebert DD, Lehr D, Smit F, Riper H, Berking M, et al. Effect of a Web-based guided self-help intervention for prevention of major depression in adults with subthreshold depression: a randomized clinical trial. JAMA 2016 May 3;315(17):1854-1863. [doi: 10.1001/jama.2016.4326] [Medline: 27139058]
2. Sander L, Rausch L, Baumeister H. Effectiveness of Internet-based interventions for the prevention of mental disorders: a systematic review and meta-analysis. JMIR Ment Health 2016;3(3):e38 [FREE Full text] [doi: 10.2196/mental.6061] [Medline: 27535468]
3. Andersson G, Titov N. Advantages and limitations of Internet-based interventions for common mental disorders. World Psychiatry 2014 Feb;13(1):4-11 [FREE Full text] [doi: 10.1002/wps.20083] [Medline: 24497236]
4. Donkin L, Glozier N. Motivators and motivations to persist with online psychological interventions: a qualitative study of treatment completers. J Med Internet Res 2012;14(3):e91 [FREE Full text] [doi: 10.2196/jmir.2100] [Medline: 22743581]
5. 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]
6. 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]
7. Mâsse LC, Watts AW, Barr SI, Tu AW, Panagiotopoulos C, Geller J, et al. Individual and household predictors of adolescents' adherence to a Web-based intervention. Ann Behav Med 2015 Jun;49(3):371-383. [doi: 10.1007/s12160-014-9658-z] [Medline: 25270826]
8. Beatty L, Binnion C. A systematic review of predictors of, and reasons for, adherence to online psychological interventions. Int J Behav Med 2016 Dec;23(6):776-794. [doi: 10.1007/s12529-016-9556-9] [Medline: 26957109]
9. Karyotaki E, Kleiboer A, Smit F, Turner DT, Pastor AM, Andersson G, et al. Predictors of treatment dropout in self-guided web-based interventions for depression: an 'individual patient data' meta-analysis. Psychol Med 2015 Oct;45(13):2717-2726. [doi: 10.1017/S0033291715000665] [Medline: 25881626]
10. Kelders SM, Bohlmeijer ET, Van Gemert-Pijnen JE. Participants, usage, and use patterns of a web-based intervention for the prevention of depression within a randomized controlled trial. J Med Internet Res 2013;15(8):e172 [FREE Full text] [doi: 10.2196/jmir.2258] [Medline: 23963284]
11. 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]
12. Zarski A, Lehr D, Berking M, Riper H, Cuijpers P, Ebert DD. Adherence to Internet-based mobile-supported stress management: a pooled analysis of individual participant data from three randomized controlled trials. J Med Internet Res 2016;18(6):e146 [FREE Full text] [doi: 10.2196/jmir.4493] [Medline: 27357528]
13. 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]
14. 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]
15. Coa K, Patrick H. Baseline motivation type as a predictor of dropout in a healthy eating text messaging program. JMIR Mhealth Uhealth 2016 Sep 29;4(3):e114 [FREE Full text] [doi: 10.2196/mhealth.5992] [Medline: 27688034]
16. Webber KH, Tate DF, Ward DS, Bowling JM. Motivation and its relationship to adherence to self-monitoring and weight loss in a 16-week Internet behavioral weight loss intervention. J Nutr Educ Behav 2010;42(3):161-167. [doi: 10.1016/j.jneb.2009.03.001] [Medline: 20138583]
17. Schwarzer R. Modeling health behavior change: how to predict and modify the adoption and maintenance of health behaviors. Applied Psychology 2008 Jan;57(1):1-29. [doi: 10.1111/j.1464-0597.2007.00325.x]
18. Sheeran P. Intention-behavior relations: a conceptual and empirical review. Eur Rev Soc Psychol 2002 Jan;12(1):1-36. [doi: 10.1080/14792772143000003]
19. Schwarzer R, Luszczynska A. Health action process approach. In: Connor P, editor. Predicting Health Behavior. 3rd Edition. Maidenhead: McGraw Hill Open University Press; 2015:252-278.
20. Schwarzer R, Lippke S, Luszczynska A. Mechanisms of health behavior change in persons with chronic illness or disability: the health action process approach (HAPA). Rehabil Psychol 2011 Aug;56(3):161-170. [doi: 10.1037/a0024509] [Medline: 21767036]
21. Schüz B, Sniehotta FF, Mallach N, Wiedemann AU, Schwarzer R. Predicting transitions from preintentional, intentional and actional stages of change. Health Educ Res 2009 Feb;24(1):64-75. [doi: 10.1093/her/cym092] [Medline: 18245046]
22. Lippke S, Ziegelmann JP, Schwarzer R. Stage-specific adoption and maintenance of physical activity: testing a three-stage model. Psychol Sport Exerc 2005 Sep;6(5):585-603. [doi: 10.1016/j.psychsport.2004.11.002]

23. Renner B, Schwarzer R. The motivation to eat a healthy diet? How intenders and nonintenders differ in terms of risk perception, outcome expectancies, self-efficacy, and nutrition behavior. Polish Psychol Bull 2005;36(1):15. [doi: 10.1016/S0065-2601(06)38002-1]
24. Reuter T, Ziegelmann JP, Lippke S, Schwarzer R. Long-term relations between intentions, planning, and exercise: a 3-year longitudinal study after orthopedic rehabilitation. Rehabil Psychol 2009 Nov;54(4):363-371. [doi: 10.1037/a0017830] [Medline: 19929117]
25. Lippke S, Plotnikoff RC. Testing two principles of the health action process approach in individuals with type 2 diabetes. Health Psychol 2014 Jan;33(1):77-84. [doi: 10.1037/a0030182] [Medline: 23088172]
26. Luszczynska A, Schwarzer R. Planning and self-efficacy in the adoption and maintenance of breast self-examination: a longitudinal study on self-regulatory cognitions. Psychol Heal 2003 Jan;18(1):93-108. [doi: 10.1080/0887044021000019358]
27. Schwarzer R, Luszczynska A. How to overcome health-compromising behaviors. Eur Psychol 2008 Jan;13(2):141-151. [doi: 10.1027/1016-9040.13.2.141]
28. Sniehotta FF, Scholz U, Schwarzer R. Bridging the intention-behaviour gap: planning, self-efficacy, and action control in the adoption and maintenance of physical exercise. Psychol Heal 2005 Apr;20(2):143-160. [doi: 10.1080/08870440512331317670]
29. 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]
30. Schwarzer R, Luszczynska A, Ziegelmann JP, Scholz U, Lippke S. Social-cognitive predictors of physical exercise adherence: three longitudinal studies in rehabilitation. Health Psychol 2008 Jan;27(1S):S54-S63. [doi: 10.1037/0278-6133.27.1(Suppl.).S54] [Medline: 18248106]
31. Lippke S, Ziegelmann J, Schwarzer R. Behavioral intentions and action plans promote physical exercise: a longitudinal study with orthopedic rehabilitation patients. J Sport Exerc Psychol 2004;26(3):483.
32. Luszczynska A, Sutton S. Physical activity after cardiac rehabilitation: evidence that different types of self-efficacy are important in maintainers and relapsers. Rehabil Psychol 2006;51(4):314-321. [doi: 10.1037/0090-5550.51.4.314]
33. Scholz U, Sniehotta F, Schwarzer R. Predicting physical exercise in cardiac rehabilitation: the role of phase-specific self-efficacy beliefs. J Sport Exerc Psychol 2005;27:151.
34. Gollwitzer P, Sheeran P. Implementation intentions and goal achievement: a meta-analysis of effects and processes. Adv Exp Soc Psychol 2008;38:119.
35. Luszczynska A, Sobczyk A, Abraham C. Planning to lose weight: randomized controlled trial of an implementation intention prompt to enhance weight reduction among overweight and obese women. Health Psychol 2007 Jul;26(4):507-512. [doi: 10.1037/0278-6133.26.4.507] [Medline: 17605571]
36. Scholz U, Schüz B, Ziegelmann JP, Lippke S, Schwarzer R. Beyond behavioural intentions: planning mediates between intentions and physical activity. Br J Health Psychol 2008 Sep;13(Pt 3):479-494. [doi: 10.1348/135910707X216062] [Medline: 17553212]
37. Sniehotta FF, Schwarzer R, Scholz U, Schüz B. Action planning and coping planning for long-term lifestyle change: theory and assessment. Eur J Soc Psychol 2005 Jul;35(4):565-576. [doi: 10.1002/ejsp.258] [Medline: 25855820]
38. Scholz U, Nagy G, Göhner W, Luszczynska A, Kliegel M. Changes in self-regulatory cognitions as predictors of changes in smoking and nutrition behaviour. Psychol Health 2009 Jun;24(5):545-561. [doi: 10.1080/08870440801902519] [Medline: 20205011]
39. Buntrock C, Ebert D, Lehr D, Riper H, Smit F, Cuijpers P, et al. Effectiveness of a web-based cognitive behavioural intervention for subthreshold depression: pragmatic randomised controlled trial. Psychother Psychosom 2015 Sep;84(6):348-358. [doi: 10.1159/000438673] [Medline: 26398885]
40. Beck A, Steer A, Brown G. BDI-II: Beck Depression Inventory Manual. 2nd edition. San Antonio: Psychological Corporation; 1996.
41. Radloff LS. A self-report depression scale for research in the general population. Appl Psychol Meas 1977;1(3):401.
42. Hautzinger M, Bailer M, Hofmeister D, Keller F. Allgemeine Depressions Skala (ADS) Manual. Boston: Hogrefe Publishers; 2012.
43. Wittchen H, Zaudig M, Fydrich T. Strukturiertes Klinisches Interview fur DSM-IV. Göttingen: Hogrefe; 1997.
44. Rohde P, Lewinsohn PM, Seeley JR. Comparability of telephone and face-to-face interviews in assessing axis I and II disorders. Am J Psychiatry 1997 Nov;154(11):1593-1598. [Medline: 9356570]
45. Allen K, Cull A, Sharpe M. Diagnosing major depression in medical outpatients: acceptability of telephone interviews. J Psychosom Res 2003 Oct;55(4):385-387. [Medline: 14507551]
46. Buntrock C, Ebert D, Lehr D, Cuijpers P, Riper H, Smit F, et al. Evaluating the efficacy and cost-effectiveness of web-based indicated prevention of major depression: design of a randomised controlled trial. BMC Psychiatry BioMed Central 2014;14:25. [doi: 10.1186/1471-244X-14-25]
47. Ebert DD, Lehr D, Heber E, Riper H, Cuijpers P, Berking M. Internet- and mobile-based stress management for employees with adherence-focused guidance: efficacy and mechanism of change. Scand J Work Environ Health 2016 Sep 01;42(5):382-394. [doi: 10.5271/sjweh.3573] [Medline: 27249161]

48. 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]
49. Meyer TD, Hautzinger M. Allgemeine Depressions-Skala (ADS). Diagnostica 2001 Oct;47(4):208-215. [doi: 10.1026//0012-1924.47.4.208]
50. Rush A, Trivedi M, Ibrahim H, Carmody T, Arnow B, Klein D, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol Psychiatry 2003 Sep 01;54(5):573-583. [Medline: 12946886]
51. Hamilton M. Development of a rating sale for primary depressive illness. Br J Soc Clin Psychol 1967;6:278-296.
52. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960 Feb;23:56-62 [FREE Full text] [Medline: 14399272]
53. Yuan KH, Bentler PM. Normal theory based test statistics in structural equation modelling. Br J Math Stat Psychol 1998 Nov;51 (Pt 2):289-309. [Medline: 9854947]
54. Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. Methods Psychological Res Online 2003;8(2):23-74.
55. Little T, Cunningham W, Shahar G, Widaman K. To parcel or not to parcel; exploring the question, weighing the merits. Struct Equ Model 2002;9(2):173. [doi: 10.1207/S15328007SEM0902]
56. Bentler PM. Comparative fit indexes in structural models. Psychol Bull 1990 Mar;107(2):238-246. [Medline: 2320703]
57. Preacher KJ, Kelley K. Effect size measures for mediation models: quantitative strategies for communicating indirect effects. Psychol Methods 2011 Jun;16(2):93-115. [doi: 10.1037/a0022658] [Medline: 21500915]
58. Lippke S, Ziegelmann J, Schwarzer R. Initiation and maintenance of physical exercise: stage-specific effects of a planning intervention. Res. in Sports Med 2004 Jul;12(3):221-240. [doi: 10.1080/15438620490497567]
59. Luszczynska A. An implementation intentions intervention, the use of a planning strategy, and physical activity after myocardial infarction. Soc Sci Med 2006 Feb;62(4):900-908. [doi: 10.1016/j.socscimed.2005.06.043] [Medline: 16095786]
60. Sniehotta FF, Scholz U, Schwarzer R. Action plans and coping plans for physical exercise: a longitudinal intervention study in cardiac rehabilitation. Br J Health Psychol 2006 Feb;11(Pt 1):23-37. [doi: 10.1348/135910705X43804] [Medline: 16480553]
61. Ziegelmann JP, Lippke S, Schwarzer R. Adoption and maintenance of physical activity: planning interventions in young, middle-aged, and older adults. Psychol Health 2006;21(2):145-163. [doi: 10.1080/1476832050018891] [Medline: 21985115]
62. Hattar A, Pal S, Hagger MS. Predicting physical activity-related outcomes in overweight and obese adults: a health action process approach. Appl Psychol Health Well Being 2016 Mar;8(1):127-151. [doi: 10.1111/aphw.12065] [Medline: 26970113]
63. Krämer LV, Helmes AW, Seelig H, Fuchs R, Bengel J. Correlates of reduced exercise behaviour in depression: the role of motivational and volitional deficits. Psychol Health 2014;29(10):1206-1225. [doi: 10.1080/08870446.2014.918978] [Medline: 24785393]
64. Parschau L, Richert J, Koring M, Ernsting A, Lippke S, Schwarzer R. Changes in social-cognitive variables are associated with stage transitions in physical activity. Health Educ Res 2012 Feb;27(1):129-140. [doi: 10.1093/her/cyr085] [Medline: 21890843]
65. 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]
66. 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]
67. 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]
68. Kenter RMF, Cuijpers P, Beekman A, van Straten A. Effectiveness of a Web-based guided self-help intervention for outpatients with a depressive disorder: short-term results from a randomized controlled trial. J Med Internet Res 2016 Mar 31;18(3):e80 [FREE Full text] [doi: 10.2196/jmir.4861] [Medline: 27032449]
69. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JEWC. 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]
70. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
71. Habibović M, Cuijpers P, Alings M, van der Voort P, Theuns D, Bouwels L, et al. Attrition and adherence in a WEB-Based Distress Management Program for Implantable Cardioverter defibrillator Patients (WEBCARE): randomized controlled trial. J Med Internet Res 2014;16(2):e52 [FREE Full text] [doi: 10.2196/jmir.2809] [Medline: 24583632]

Abbreviations

AES: Advanced Encryption Standard

CES-D: Center for Epidemiological Studies Depression Scale
CFI: comparative fit index
eCoach: electronic coach
HAPA: health action process approach
HRSD24: Hamilton Rating Scale for Depression
QIDS-CR16: Quick Inventory of Depressive Symptomatology–Clinician Rating
RCT: randomized controlled trial
RMSEA: root mean square error approximation
SRMR: standardized root mean square residual
TLI: Tucker-Lewis Index

Edited by G Eysenbach; submitted 26.08.17; peer-reviewed by L Krämer, I Beintner, R Kok; comments to author 13.09.17; revised version received 27.09.17; accepted 28.09.17; published 11.01.18.

Please cite as:
Zarski AC, Berking M, Reis D, Lehr D, Buntrock C, Schwarzer R, Ebert DD
Turning Good Intentions Into Actions by Using the Health Action Process Approach to Predict Adherence to Internet-Based Depression Prevention: Secondary Analysis of a Randomized Controlled Trial
J Med Internet Res 2018;20(1):e9
URL: https://www.jmir.org/2018/1/e9/
doi:10.2196/jmir.8814
PMID:29326097

Original Paper

A Tailored Web-Based Intervention to Improve Parenting Risk and Protective Factors for Adolescent Depression and Anxiety Problems: Postintervention Findings From a Randomized Controlled Trial

Marie Bee Hui Yap[1,2], MPsych (Clin), PhD; Shireen Mahtani[1], DPsych (Clin); Ronald M Rapee[3], PhD; Claire Nicolas[1], BSocSc, GradDipPsych; Katherine A Lawrence[1], DPsych (Clin); Andrew Mackinnon[2,4], PhD; Anthony F Jorm[2], PhD, DSc

[1]Monash Institute of Cognitive and Clinical Neurosciences, School of Psychological Sciences, Monash University, Clayton, Australia
[2]Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Australia
[3]Centre for Emotional Health, Macquarie University, New South Wales, Australia
[4]Black Dog Institute, University of New South Wales, Sydney, New South Wales, Australia

Corresponding Author:
Marie Bee Hui Yap, MPsych (Clin), PhD
Monash Institute of Cognitive and Clinical Neurosciences
School of Psychological Sciences
Monash University
18 Innovation Walk
Clayton, 3800
Australia
Phone: 61 399050723
Email: marie.yap@monash.edu

Abstract

Background: Depression and anxiety disorders in young people are a global health concern. Parents have an important role in reducing the risk of these disorders, but cost-effective, evidence-based interventions for parents that can be widely disseminated are lacking.

Objective: This study aimed to examine the postintervention effects of the Partners in Parenting (PiP) program on parenting risk and protective factors for adolescent depression and anxiety, and on adolescent depression and anxiety symptoms.

Methods: A two-arm randomized controlled trial was conducted with 359 parent-adolescent dyads, recruited primarily through schools across Australia. Parents and adolescents were assessed at baseline and 3 months later (postintervention). Parents in the intervention condition received PiP, a tailored Web-based parenting intervention designed following Persuasive Systems Design (PSD) principles to target parenting factors associated with adolescents' risk for depression and anxiety problems. PiP comprises a tailored feedback report highlighting each parent's strengths and areas for improvement, followed by a set of interactive modules (up to nine) that is specifically recommended for the parent based on individually identified areas for improvement. Parents in the active-control condition received a standardized package of five Web-based factsheets about adolescent development and well-being. Parents in both conditions received a 5-min weekly call to encourage progress through their allocated program to completion. Both programs were delivered weekly via the trial website. The primary outcome measure at postintervention was parent-reported changes in parenting risk and protective factors, which were measured using the Parenting to Reduce Adolescent Depression and Anxiety Scale (PRADAS). Secondary outcome measures were the adolescent-report PRADAS, the parent- and child-report Short Mood and Feelings Questionnaire (depressive symptoms), and parent- and child-report Spence Children's Anxiety Scale (anxiety symptoms).

Results: Parents in the intervention condition completed a mean of 73.7% of their intended personalized PiP program. A total of 318 parents (88.6%, 318/359) and 308 adolescents (92.8%, 308/332) completed the postintervention assessment. Attrition was handled using mixed model of repeated measures analysis of variance. As hypothesized, we found a significant condition-by-time interaction on the PRADAS, with a medium effect size, Cohen *d*=0.57, 95% CI 0.34-0.79. No significant differences between conditions were found at postintervention on any of the secondary outcome measures, with adolescent depressive (parent-report only) and anxiety (both parent- and adolescent-report) symptoms decreasing significantly from baseline to postintervention in both conditions.

Conclusions: The fully automated PiP intervention showed promising short-term effects on parenting behaviors that are associated with adolescents' risk for depression and anxiety. Long-term follow-up is required to ascertain whether these effects translate into reduced adolescent depression and anxiety problems. The intervention may be useful as a low-cost universal public health program to increase parenting practices believed to benefit adolescents' mental health.

Trial Registration: Australia New Zealand Clinical Trials Registry: ACTRN12615000328572; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx? id=368274 (Archived by WebCite at http://www.webcitation.org/6qgsZ3Aqj)

(J Med Internet Res 2018;20(1):e17) doi:10.2196/jmir.9139

KEYWORDS

family; anxiety; parenting; depression; adolescent; Internet; mental health; preventive health services

Introduction

The Problem of Depression and Anxiety Disorders in Youth

Depression and anxiety disorders are among the most common mental disorders affecting 18% and 38%, respectively, of young people in the age range of 13 to 17 years [1]. Developing these disorders early in life, especially if left untreated, can increase young people's suicide risk and forecast a wide range of psychosocial and vocational impairments, resulting in deleterious long-term sequelae [2-4]. Although intervention efforts for these disorders continue to progress, a large proportion of the burden of disease is still unavertable even with optimal treatment [5]. There is, hence, an urgent need for an effective, integrated approach to prevent these disorders. As the incidence of depression and anxiety disorders peaks during adolescence, early adolescence is a particularly opportune time to target preventive efforts [6]. Fortunately, evidence to date indicates that depression and anxiety disorders in young people can be prevented [7-9].

Parents Have an Important Role in Prevention

There is now substantial robust evidence delineating risk and protective factors for adolescent anxiety and depressive disorders [10,11]. Importantly, some of these factors are within parents' control or influence and are potentially modifiable [12]. Synthesizing longitudinal, retrospective, and cross-sectional evidence, a recent review identified a sound evidence base for three protective parental factors for depression (warmth, autonomy granting, and monitoring) and one for anxiety (warmth). Three risk factors for both outcomes were also identified: interparental conflict, overinvolvement, and aversiveness [13]. However, despite this evidence base, parents' knowledge about their role in reducing their adolescent's risk of depression is less than optimal [14], highlighting a need to equip parents in the general population with evidence-based preventive resources.

Existing Preventive Parenting Interventions

Preventive parenting programs have been developed to capitalize on parents' influence on their child's development and mental health based on the underlying assumption that changing parenting (mediators) will in turn change a child's risk for depression and anxiety [12]. A recent systematic review and meta-analysis found that preventive interventions primarily targeting parents (ie, most of the intervention is with the parent, as opposed to the child or involving the whole family) have beneficial effects on the child's internalizing (depression and anxiety) outcomes lasting up to 11 years post intervention [9]. This contrasts with the evidence for preventive interventions targeting young people directly, which have observed intervention effects lasting less than 2 years [15,16]. Although this evidence base highlights the remarkable promise of preventive parenting interventions, only 3 of the 51 parenting interventions included in the review were designed for parents of adolescents [9]. Moreover, many parenting programs are not well used even when available because of barriers such as scheduling difficulties and privacy concerns [17].

Preventive parenting interventions can be *universal* (ie, delivered to all parents regardless of risk), *selective* (targeting parents whose children have known risk factors), or *indicated* (targeting parents whose children show signs or symptoms of emerging disorders) [18]. Universal programs tend to have a smaller effect than selective or indicated programs at the level of the individual [16]. However, they can have a great public health impact because they reach a larger proportion of the population [19] and have the potential to shift the population mean levels of depression and anxiety symptoms [19]. Notably, in the aforementioned review of preventive parenting interventions [9], there was no evidence that type of prevention (universal, selective, or indicated) moderated intervention effects. As highlighted in the Institute of Medicine report [18], universal interventions are advantageous when they are effective and acceptable, have a low cost per individual, and carry a low risk of harm. When trying to engage parents in prevention of mental health problems in their child, universal approaches can increase acceptability because they minimize the perceived stigma that some parents fear would be attached to themselves as "bad" parents, or to their child as having problems needing intervention [20]. Hence, a universal program should be considered an integral component in a public health approach to empower parents for their role in prevention of youth depression and anxiety disorders [9].

Potential of a Web-Based Parenting Intervention

Web-based media are a promising mode of delivering universal prevention programs because of their scalability and likely cost-effectiveness [18]. Universal programs are also well-suited to Web-based delivery because they usually involve a lower intensity of intervention (eg, require little or no contact with trained professionals), hence, reducing the cost of population-wide dissemination. Given the increasing reach of the Internet [21], Web-based media have been recommended as one effective way to increase participation in preventive interventions [22]. Web-based universal parenting programs also have the potential to overcome the aforementioned barriers of existing face-to-face programs because of their anonymity, flexibility, and accessibility. The Internet has become a popular source of information on parenting and child mental health among parents [2,23], and a recent survey found that the idea of a tailored online parenting program for parents of adolescents was viewed favorably [24]. Moreover, implementation fidelity is guaranteed by the computerized delivery of a well-designed and well-maintained program [25]. Despite these potential benefits, a recent systematic review [9] failed to identify any evidence-based, tailored Web-based parenting intervention designed to prevent adolescent depression and anxiety disorders. The potential of online prevention programs targeting parents of adolescents remains as yet, largely untapped [26], but these programs would comprise a promising public health approach to preventing adolescent depression and anxiety that is potentially lower in cost per individual compared with existing programs.

The Partners in Parenting (PiP) program is a tailored Web-based parenting intervention to prevent adolescent depression and anxiety problems (see [27] for more details). Its content is derived from Parenting Guidelines [28], which were developed through a rigorous two-stage process involving a systematic review of parenting risk and protective factors associated with adolescent depression and anxiety [12]; and a Delphi study of international expert consensus about parenting strategies that are important for reducing risk for adolescent depression and anxiety disorders [29]. The program development process was aligned with a consumer-engagement approach [30] by involving parents of adolescents in reference group workshops and obtaining input from adolescents through focus group consultations. Design of the Web-based components of PiP were guided by the Persuasive Systems Design (PSD) model that proposes to purposefully use technology to influence behavior change [31] and has been found to influence adherence to Web-based interventions [32]. For example, following the PSD tailoring principle, the program's automated tailoring feature screens each parent on a wide range of parenting factors known to influence risk for adolescent depression and anxiety. This identifies areas for improvement to target in each parent's personalized intervention. This tailoring feature increases the perceived relevance of the program for each parent [33], and potentially its effects [32], without requiring the costly involvement of trained professionals, hence increasing potential for scalability and sustainability [33]. As a preventive parenting intervention, PiP is designed to increase parental protective factors and decrease parental risk factors associated with adolescent depression and anxiety. The change in parenting factors (proximal outcome and direct target of the intervention) is in turn expected to reduce adolescent risk for depression and anxiety problems in the long term.

The primary aim of this study was to evaluate the effects of PiP compared with an active-control condition (educational factsheets on adolescent development and well-being) in a randomized controlled trial (RCT). Specifically, we hypothesized that compared with the control group, parents who received PiP will show (1) greater improvements in parenting risk and protective factors from baseline to postintervention (primary outcome) using the Parenting to Reduce Adolescent Depression and Anxiety Scale (PRADAS; [34]), a criterion-referenced measure of parenting against the Parenting Guidelines [28], and (2) greater reductions in adolescent depression and anxiety symptoms and greater improvements in adolescent-report parenting from baseline to postintervention (secondary outcomes).

Although the ultimate aim of the PiP is the prevention of adolescent depression and anxiety problems, this parenting intervention is posited to result in adolescent benefits indirectly through its effects on parenting. The RCT includes a postintervention assessment (focus of this paper), where the primary outcome of interest is the intervention's proposed mechanism of change: parenting risk and protective factors. Adolescent depressive and anxiety symptoms are secondary outcomes at postintervention because we expect the intervention's effects on these outcomes to take time, operating through changes in parenting. However, beyond the scope of this paper, the RCT also includes a 12-month follow-up assessment, which is currently being undertaken, when the primary outcomes of interest will be adolescent depressive and anxiety symptoms.

Methods

Trial Design

This study was a parallel-group superiority RCT with parent-adolescent dyads randomly allocated in a 1:1 ratio to one of the two conditions: (1) PiP or (2) educational factsheets (control intervention). The trial was prospectively registered with the Australian New Zealand Clinical Trials Registry (ACTRN12615000328572; see Multimedia Appendix 1 for Consolidated Standards of Reporting Trials [CONSORT-EHEALTH checklist]).

Setting, Participants, and Eligibility Criteria

From August 2015 to September 2016, 359 parent-adolescent dyads were recruited primarily via government, Catholic, and independent schools across the state of Victoria in Australia. Schools were contacted by email and phone to request that recruitment flyers (hard copy or electronic) were distributed to parents of students in Years 7 to 10 (aged 12-15 years). Other means of recruitment included disseminating advertisement flyers via social media, online networks, and through mental health organizations (eg, *beyondblue* and Mental Health First Aid Australia). Interested and eligible parents were invited to register via the dedicated trial website. To be eligible, parents had to have a target child in the age range of 12 to 15 years, regular access to the Internet and an email account, and reside in Australia. Computer or Internet literacy was an implicit eligibility criterion. Parents were asked to provide consent and contact details for their child to participate in the trial but could still participate if their adolescent declined participation. For each family, only one parent and one child could be included in the trial (see Multimedia Appendix 2 for participant informed consent documentation). Participants were not excluded if the adolescent scored in the clinically elevated range (as determined by published clinical cut-off scores) on either the depressive or anxiety symptom measures (either parent- or child-report) at baseline.

Figure 1 shows the study design and flow of participants. The study was primarily conducted online via automated emails to parents and a dedicated RCT website through which parents received their allocated intervention and parent and adolescent participants completed their study assessments. Adolescent assessments were completed online with the assistance of a research officer over the phone. Parents received an automated email inviting them to complete their online assessment as soon as their child had submitted their corresponding assessment responses or declined to participate.

Interventions

In addition to receiving their Web-based program (described below), all parent participants received a weekly phone call from a researcher, starting 7 days after completing their baseline survey and every week thereafter until they had completed their allocated intervention. In the intervention group, the total number of weekly calls was designed to match the number of modules in each parent's tailored parenting program. However, if they had less than five modules, they still received five calls (to match the number of calls received by the control group). Research staff were trained to make these calls following a standard protocol (ie, a flowchart of prompts and appropriate responses, eg, “Did you [complete your module or read your factsheet] this week?” and “Did you try to put into practice or apply any of the information you read?”) and did not provide individual advice or therapy. These calls were intended to address any study-related questions or troubleshoot technical issues that arose, encourage parents to progress through their allocated intervention each week till completion, and enhance parents' engagement.

The Partners in Parenting Intervention

PiP [27] is a Web-based parenting program that is part of the broader Parenting Strategies research translation online platform [35]. The programming of the intervention was first completed in July 2015 and was not modified throughout the trial.

On the basis of their responses to a self-assessment parenting scale (the PRADAS [34]), parents in the intervention condition received an individually tailored feedback report that highlighted areas where they were doing well (ie, concordant with the Parenting Guidelines [28]) and areas where they could improve (ie, not concordant with the Guidelines). They were then given access to the Web-based modules (up to nine) to support them in making changes to identified areas for improvement [27]. Specific modules were recommended to parents based on their responses to the PRADAS at baseline. Table 1 shows the alignment of topics across the Guidelines, PRADAS, feedback report, and modules (screenshots are available in Multimedia Appendix 3).

Feedback messages in PiP are brief, with practical strategies provided in dot point form and are designed to motivate behavior change [27]. Parents viewed their feedback report on the website immediately after submitting their online baseline assessment and being randomly allocated to the intervention condition. They were also emailed a copy of their feedback report, the Parenting Guidelines, and instructions on accessing the interactive Web-based parenting program with their recommended modules.

Upon logging in to their parenting program, parents were presented with their recommended modules, as well as other available modules. They could further tailor their program at this stage by deselecting recommended modules and/or selecting additional modules. They then confirmed their selection and commenced their personalized program. The nine modules comprising the PiP intervention were derived from topics covered in the Parenting Guidelines (see Table 1). Modules include illustrations, audio clips, vignettes, interactive activities, goal-setting exercises, and an end-of-module quiz with immediate feedback to consolidate learning. Each module is designed to help parents make changes to their parenting so as to become more concordant with the Guidelines. Each module takes about 15 to 25 min to complete. Parents were directed to their first selected module immediately after they had completed their baseline assessment and received their feedback report. Thereafter, parents were notified via weekly automated emails as their next module was made available for them through their personalized dashboard until they had completed their whole program. One module is made available for parents every 7 days, in a set order, regardless of whether they had completed preceding modules. After completing their program, parents had unlimited access to all PiP modules for the duration of the RCT (up to 3 years for parents who registered early in the recruitment phase of the 3-year RCT, ie, August 2015).

Educational Factsheets (Control Intervention)

Parents in the control condition were provided with a standardized package of educational materials about adolescent development and mental health via the trial website. Each week for 5 weeks, parents received an automated email inviting them to access their factsheet for that week (to match the expected mean number of modules received by the intervention group). To mirror the experience of intervention group parents who accessed each module on the trial website, control group parents accessed each factsheet by logging in to their dashboard on the website. The factsheets provide general information to parents as opposed to tailored, actionable strategies and were designed to represent a selection of resources that are available to parents as part of the current health promotion approach for adolescent well-being. The materials were adapted from credible existing resources provided on the Raising Children Network website [36]. The topics of the five factsheets were (1) Teen development: an overview, (2) The teenager's developing brain, (3) The teenager's changing body, (4) Resilience, and (5) Happy teenagers and teenage well-being. Parents had access to these factsheets for the duration of the RCT.

Figure 1. Participant flow diagram. ITT=Intention-to-treat analyses. Parent or dyad remain enrolled in study unless indicated that they had opted out. aIncludes complete dyads opted out; badolescent opted out of study at 3-month follow-up; cof these, 2 adolescent participants had opted out before completing adolescent baseline survey and, one adolescent's 3-month follow-up was also missed because of a technical error; dadolescent participant(s) opted out before completing adolescent baseline survey; eadolescent participant opted out of completing 3-month follow-up survey; and fadolescent participant completed 3-month follow-up.

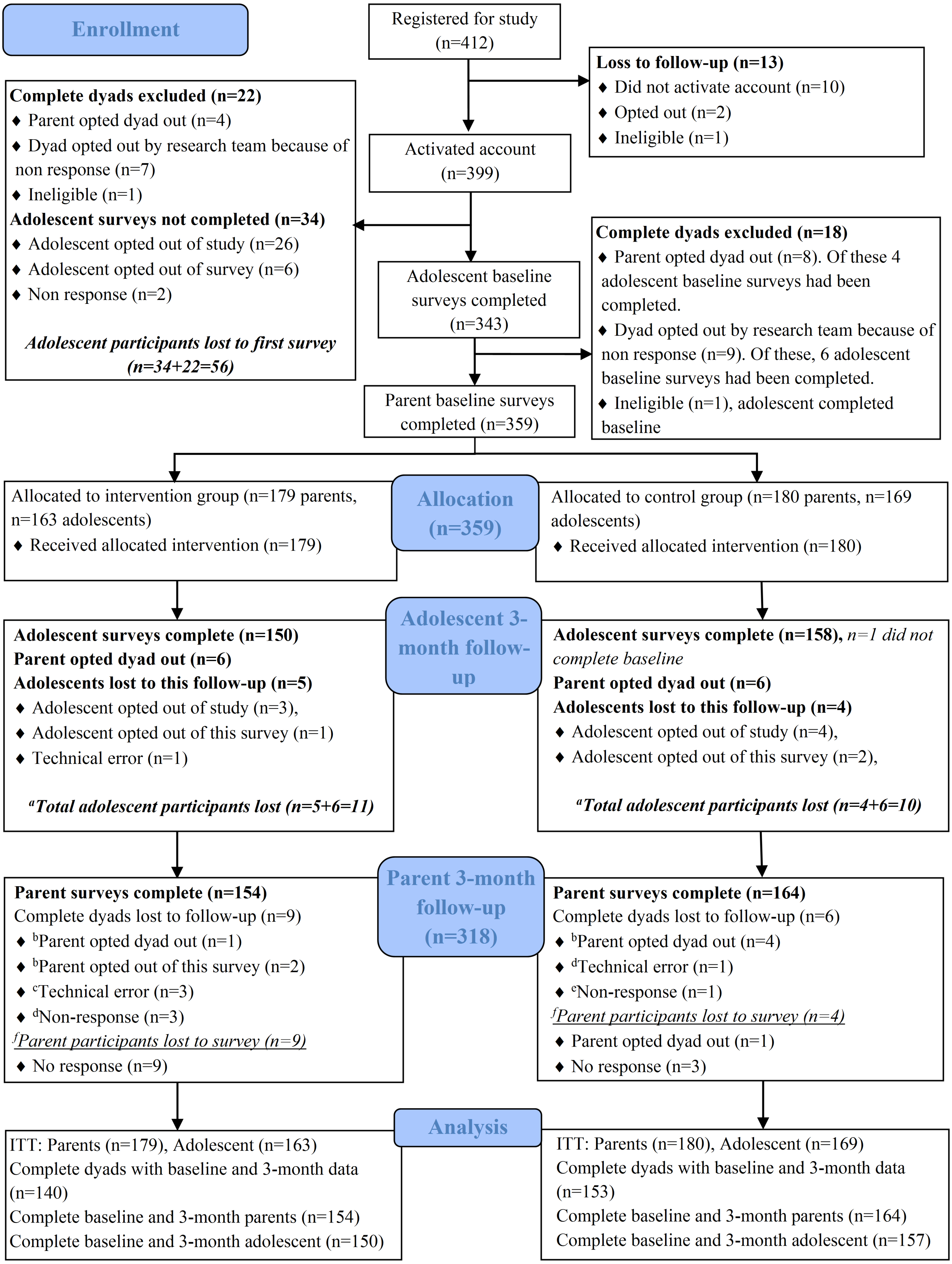


Table 1. Guidelines topics, corresponding sections of the parenting scale (Parenting to Reduce Adolescent Depression and Anxiety Scale, PRADAS) and personalized feedback report, title of interactive modules, and outline of content.

Guidelines subheading	Corresponding section of the PRADAS[a] and feedback report	Title of interactive module	Outline of content
You can reduce your child's risk of developing depression and clinical anxiety	Not applicable (NA). Not included in the PRADAS or feedback report	NA. No module on this topic	Psychoeducation about the role of parents in the prevention of adolescent depression and anxiety
Establish and maintain a good relationship with your teenager	Your relationship with your teenager	Connect	Acknowledges the challenge of connecting with adolescent children and provides specific tips on how to do this
Be involved and support increasing autonomy	Your involvement in your teenager's life	Nurture roots and inspire wings	Helps parents establish the important balance between staying involved and interested in their adolescent's life, while encouraging increasing age-appropriate autonomy
Encourage supportive relationships	Your teenager's relationships with others	Good friends, supportive relationships	Provides strategies for parents to support their adolescent's social skills development
Establish family rules and consequences	Your family rules	Raising good kids into great adults: establishing family rules	Highlights the importance of consistent and clear boundaries for adolescent behaviors and provides specific strategies to establish these
Minimize conflict in the home	Your home environment	Calm versus conflict	Addresses the need for adaptive conflict management between parents and between parent and adolescent and provides specific strategies to do these
Encourage good health habits	Health habits	Good health habits for good mental health	Provides strategies to help parents encourage good health habits in their adolescent, including a healthy diet, physical activity, good sleep habits, and abstinence from alcohol and drugs
Help your teenager to deal with problems	Dealing with problems in your teenager's life	Partners in problem solving	Provides strategies for parents to help their adolescent develop good problem-solving and stress management skills
Help your teenager to deal with anxiety	Coping with anxiety	From surviving to thriving: helping your teenager deal with anxiety	Provides strategies for parents to help their adolescent manage their everyday anxiety
Encourage professional help seeking when needed	Getting help when needed	When things aren't okay: getting professional help	Helps parents understand what depression and anxiety problems can look like in adolescents, and what they can do if their adolescent is or becomes unwell
Don't blame yourself	Don't blame yourself (included for all parents in feedback report only)	NA. No module on this topic	Aims to dispel guilt or self-blame in parents

[a] PRADAS: Parenting to Reduce Adolescent Depression and Anxiety Scale.

Outcomes

Parents and adolescents were asked to complete online assessments comprising the following measures at baseline and 3 months later. The latter time point was designated as the postintervention assessment based on the expectation that most parents would have completed their intervention (maximum of nine modules, one per week) by then.

Primary Outcome Measure: Parenting to Reduce Adolescent Depression and Anxiety Scale (PRADAS)

The PRADAS [34] is a criterion-referenced measure assessing parents' current parenting behaviors against specific recommendations in the Parenting Guidelines (the "criterion"; [28]). Newly developed for use in this study, the original version consisted of 79 questions assessing the nine domains of parenting addressed in the Parenting Guidelines. Following validation analyses [34], the scale was revised to include 73 of the original 79 items after dropping one 6-item subscale ("Your teenager's relationships with others"). The final scale assesses each parent's total concordance with the Parenting Guidelines across eight domains of parenting assessed using eight subscales (8-12 items per subscale). Each item corresponds to a parenting recommendation in the Guidelines and involves either a 4-point Likert-type frequency scale (for specific parenting behaviors such as showing affection to their child; eg, never, rarely, sometimes, or often) or a likelihood scale (for hypothetical scenarios such as noticing a persistent change in their adolescent's behavior; eg, very unlikely, unlikely, likely, or very likely). See Multimedia Appendix 3 for sample items.

For each item, certain responses were prespecified by the authors [34] as indicating concordance with the Guidelines (scored as 1), with the remaining responses for the item deemed "nonconcordant" (scored as 0). Scores for all 73 items were summed to yield the total concordance score, which can range from 0 to 73. The PRADAS total score demonstrated good reliability (including coefficient of agreement and test-retest reliability) and convergent validity in an Australian validation sample comprising 711 parents, which included data from the current RCT [34]. Test-retest reliability for the PRADAS total score was r=.76, P<.001 for a smaller subsample of 175 parents based on data collected 1 month apart [34]. The coefficient of agreement for the PRADAS total score in the current sample was 0.97 at baseline and 0.96 at 3-month follow-up assessment.

Secondary Outcome Measure #1: Short Mood and Feelings Questionnaire (SMFQ)

The Short Moods and Feelings Questionnaire (SMFQ) is the brief 13-item version of the Mood and Feelings Questionnaire [37], which assesses depressive symptoms in children and adolescents using child-report (SMFQ-C) or parent-report (SMFQ-P). Respondents indicate the frequency of depressive symptoms in the past 2 weeks by rating each item on a 3-point scale, with 0=not true, 1=sometimes true, and 2=true. Both parents and adolescents reported on adolescent depressive symptoms in this study; Cronbach alphas were .91 and .90, respectively, at baseline and .91 and .92 at postintervention.

Secondary Outcome Measure #2: Spence Children's Anxiety Scale (SCAS)

Adolescent anxiety symptoms were assessed using the *Spence Children's Anxiety Scale-Child version* (SCAS-C) [38] and SCAS for parents (SCAS-P; [39]). Both versions have 39 items assessing six domains of anxiety in children: separation anxiety, social phobia, obsessive-compulsive disorder, panic disorder or agoraphobia, generalized anxiety, and fear of physical injury. Respondents rate the frequency of anxiety symptoms on a 4-point scale, from 0=never to 3=always. Item responses are summed to form a total anxiety score, which has demonstrated acceptable-to-high reliability in children in the age range of 8 to 14 years. Cronbach alphas were .91 and .92 at baseline and .94 and .95 at postintervention for parent and adolescent versions, respectively.

Secondary Outcome Measure #3: Parenting to Reduce Adolescent Depression and Anxiety Scale-Adolescent (PRADAS-A) Report

An adolescent-report version of the PRADAS, the PRADAS-A, was developed and validated by the research team, also as a criterion-referenced measure (Cardamone-Breen, unpublished data, 2017). The original PRADAS-A has 47 items and assesses a subset of parenting behaviors across the same nine domains assessed in the PRADAS. Following validation analyses, the same "Relationships with others" subscale (four items) was dropped. The final PRADAS-A comprises 43 items in eight subscales (2-7 items per subscale). Most items involve a 4-point Likert-type frequency response scale (ie, never, rarely, sometimes, and often) except for the last 2-item subscale "Getting help when needed," which utilized a 4-point likelihood scale (ie, very unlikely, unlikely, likely, and very likely).

Scoring of responses on each item as reflecting parenting behaviors that are concordant (scored as 1) or nonconcordant (scored as 0) with the Parenting Guidelines was predetermined by the scale authors. Scores for all 43 items were added to form the total concordance score, which can range from 0 to 43. The PRADAS-A total score demonstrated good reliability (including coefficient of agreement and test-retest reliability) and convergent validity in an Australian validation sample comprising 670 adolescents in the age range of 12 to 15 years, which included data from the current trial (Cardamone-Breen, unpublished data, 2017). Test-retest reliability for the PRADAS-A total score was r=.81, P<.001 for a smaller subsample of 160 adolescents participants based on data collected 3 months apart (Cardamone-Breen, unpublished data, 2017). The coefficient of agreement for the PRADAS-A total score in the current sample was 0.97 at both baseline and 3-month follow-up assessments.

Intervention Completion and Adherence

We defined intervention adherence following Kelders and colleagues [40] whereby percentage of individuals who adhere to the intervention=100% x [(number of participants whose observed usage equals their intended usage of the Web-based intervention)/(total number of individuals who received the intervention)]. We operationalized intervention completion as percentage of program completed=100% x [(observed usage of the Web-based intervention) or (intended usage of the Web-based intervention)]. For the PiP condition, intended usage is the total number of modules the parent had locked in to their personalized program after reviewing the program's recommendations and applying their personal preferences. Observed usage is defined as the total number of modules in their personalized program that were completed. A module was considered to be completed if it had a "closed" timestamp, there were responses recorded for its end-of-module quiz, or the selected goal for the module had been checked off as achieved by the parent. A closed timestamp for each module was stored in the database when the parent clicked on "Finish module" on the last page of the module. For the control condition, intended usage is fixed as five factsheets, whereas observed usage was the total number of factsheets with a closed timestamp. A closed timestamp for each factsheet was saved as long as the factsheet was clicked on and "opened" on the website.

Sample Size

For a repeated measures design, with one preintervention measure and two postintervention measures, using the analysis of covariance method of sample size calculation and assuming a 0.70 correlation between pre-post measurements, to detect a small effect size (Cohen d=0.20), with a power of 0.80 and a Cronbach alpha of .05, we required a total sample of 294 parent-adolescent dyads (147 dyads per group). Allowing for up to 15% attrition, we aimed to recruit 338 parent-adolescent dyads (169 dyads per group).

Randomization

The random sequence generation was automated within the trial website via software architecture, with participant assignment revealed to parents only after parent and adolescent (if applicable) baseline assessments had been completed, hence, ensuring allocation concealment. Parent-adolescent dyads were randomly assigned in a 1:1 ratio with no stratification, resulting in 179 dyads allocated to PiP and 180 dyads to the control condition.

Blinding

It was not possible to blind parents to their assignment because of the informed consent procedures. The research officers who spoke to parents for their weekly calls were also not blinded to the parent's assignment as they had to tailor the script to match the program the parent was receiving. Adolescents were not informed of their parent's assignment, so were assumed to be "blinded" to this.

Statistical Methods

All statistical analyses were completed with Statistical Package for the Social Sciences (SPSS) version 24.0 (IBM Corp) software. We conducted independent *t* tests and chi-square tests to test for differences between conditions at baseline in outcome measures and demographic data. We also compared the baseline characteristics of those who completed postintervention assessments and those who did not, to explore possible attrition biases.

We conducted all analyses on an intention-to-treat basis, using mixed model of repeated measures (MMRM), a likelihood-based approach that utilizes all available data, including those from participants who withdrew from the trial after completing the baseline assessment. MMRM produces unbiased estimates of intervention effects under the assumption that data are missing completely at random or missing at random [41].

For MMRM, intervention group (condition) and time ie, assessment wave) and the interaction between condition and time were set as fixed factors.

Post-hoc analyses were also conducted to explore whether intervention effects differed depending on adolescents' depression or anxiety symptom levels at baseline. SPSS PROCESS macros [42] were used to conduct these moderation analyses, with the predictor of condition coded as 1=PiP and 0=control, mean-centered baseline SCAS or SMFQ score as the continuous moderator, and change in outcome variables computed by subtracting baseline scores from postintervention scores. To minimize shared method variance, we relied on different informants (parent vs adolescent) for the moderator versus outcome variables. For example, to assess effects on SMFQ-P, we used SMFQ-C scores as the moderator. To minimize the number of post-hoc analyses conducted, when examining a parenting outcome according to one informant, we relied on the other informant for the moderator (one of the two symptom measures). Hence, for example, the PRADAS was the outcome variable in two moderation analyses, one using change in SMFQ-C scores as the moderator and the other using the SCAS-C.

All tests were conducted using Cronbach alpha level of .05 and 95% CIs.

Ethics and Informed Consent

This RCT was approved by the Monash University Human Research Ethics Committee, CF14/3887-2014002024. Informed consent was obtained from parent participants at registration via checkboxes on the trial website, and verbal assent from adolescent participants was obtained over the phone.

Safety Protocols

Participants were followed up by a provisional psychologist (postgraduate doctoral candidate in clinical psychology) if the parent-adolescent dyad both reported elevated symptoms in the adolescent based on predetermined cut-off scores on the SCAS and SMFQ. Follow-up actions comprised risk assessment phone calls, with adolescent participants and email notifications to parents suggesting appropriate avenues for supporting their child's mental health.

Results

Randomization and Study Attrition

A total of 359 parents and 332 adolescents completed the baseline assessment. Of these, 318 parents (88.6%, 318/359) and 308 adolescents (92.8%, 308/332; this includes one adolescent who had not completed the baseline assessment) completed the postintervention assessment. This represents attrition rates of 8.9% parents (n=16) and 6.5% adolescents (n=11) from the control group and 14.0% parents (n=25) and 8.0% adolescents (n=13) from the intervention group. Figure 1 provides further details on participant flow from enrollment to postintervention assessment organized according to the CONSORT guidelines.

Missing Data and Distributional Assumptions

Scales with a small number of items missing had these items replaced with the participant's mean response for that scale. This led to a maximum of 23% of items (3/13) being imputed for one participant, but fewer than 11% were imputed in the remaining cases. Given the low rates of missing data at the item-level per participant, all cases with complete baseline and postintervention assessments as reflected in Figure 1 were retained for analyses. Mean imputation is acceptable when data is missing for less than 5% of cases in a dataset [43].

Model residuals for the symptom outcomes (ie, parent and child SMFQ and SCAS) at pre- and postintervention were positively skewed. Log transformation addressed this deviation from normality. For ease of interpretation, we have reported findings based on the raw data in the remaining sections because findings from analyses using the raw and transformed datasets were largely similar, except where specified below (see Multimedia Appendix 4 for MMRM analyses using transformed symptom outcome variables).

Table 2. Sample characteristics at baseline by intervention condition.

Participant characteristic	Intervention (N=179)	Control (N=180)	*t* or χ^2	*P* value
Parent sex			0.7	.42
Male, n (%)	26 (14.5)	20 (11.1)		
Female, n (%)	153 (85.5)	160 (88.9)		
Parent age (years), mean (SD)	45.2 (5.26)	45.1 (5.14)	0.14	.89
Parent marital status			3.2	.36
Single, n (%)	12 (6.7)	9 (5.0)		
Married or de facto, n (%)	138 (77.1)	137 (76.1)		
Separated or divorced, n (%)	27 (15.1)	34 (18.9)		
Widowed, n (%)	2 (1.1)	0 (0)		
Child sex			0.2	.63
Male, n (%)	102 (57.0)	97 (53.9)		
Female, n (%)	77 (43.0)	83 (46.1)		
Child age, mean (SD)	13.7 (1.05)	13.7 (1.08)	−0.43	.67
Family situation			8.8	.07
Child participant lives with both parents, n (%)	131 (73.2)	122 (67.8)		
Parents separated but both involved in care of child participant, n (%)	21 (11.7)	36 (20.0)		
Parents separated with only registered parent involved in care of child participant, n (%)	16 (8.9)	14 (7.8)		
Sole parent of child participant, n (%)	10 (5.6)	4 (2.2)		
Other	1 (0.6)	4 (2.2)		
Number of children, mean (SD)	2.37 (0.94)	2.32 (1.00)	0.45	.65
Language			0.01	.98
English, n (%)	150 (83.8)	152 (84.4)		
Other, n (%)	29 (16.2)	28 (15.6)		
Parent employment			1.6	.46
Unemployed, n (%)	21 (11.7)	27 (15.0)		
Part-time, n (%)	81 (45.3)	71 (39.4)		
Full-time, n (%)	77 (43.0)	82 (45.6)		
Parent studying status			1.6	.46
Not studying, n (%)	149 (83.2)	145 (80.6)		
Studying part-time, n (%)	4 (2.2)	2 (1.1)		
Studying full-time, mean (SD), n (%)	26 (14.5)	33 (18.3)		
Parent’s highest education level, mean (SD)			4.7	.46
Year 7-12, n (%)	26 (14.5)	24 (13.3)		
Trade or apprenticeship, n (%)	2 (1.1)	4 (2.2)		
Other technical or further education (TAFE) or technical, n (%)	18 (10.1)	12 (6.7)		
Diploma, n (%)	26 (14.5)	38 (21.1)		
Bachelor degree, n (%)	63 (35.2)	56 (31.1)		
Postgraduate degree, n (%)	44 (24.6)	46 (25.6)		
Parent’s mental health diagnosis, mean (SD)			8.3	.22
None, n (%)	72 (40.2)	72 (40.0)		
Past history, n (%)	60 (33.5)	77 (43.8)		

Participant characteristic	Intervention (N=179)	Control (N=180)	t or χ^2	P value
Current diagnosis, n (%)	25 (14.0)	17 (9.4)		
Past and current diagnosis, n (%)	20 (11.2)	13 (7.2)		
Unanswered, n (%)	2 (1.1)	1 (0.6)		
Child's past mental health diagnosis, mean (SD)			10.6	.30
Depression, n (%)	3 (1.7)	0 (0)		
Any anxiety disorder, n (%)	11 (6.1)	13 (7.2)		
Autism or Asperger's syndrome, mean (SD), n (%)	4 (2.2)	5 (2.8)		
Other, n (%)	4 (2.2)	4 (2.2)		
Multiple diagnoses, mean (SD), n (%)	5 (2.8)	8 (4.4)		
No formal diagnosis, but parent concerned, n (%)	31 (17.3)	48 (26.7)		
No past diagnosis, n (%), n (%)	105 (58.7)	88 (48.9)		
Unanswered, n (%)	16 (8.9)	14 (7.8)		
Child's current mental health diagnosis			5.4	.50
Depression, n (%)	0	1 (0.6)		
Any anxiety disorder, n (%)	13 (7.3)	13.0 (7.2)		
Autism or Asperger's[a] syndrome, n (%)	3 (1.7)	4 (2.2)		
Other, n (%)	6 (3.4)	3 (1.7)		
Multiple diagnoses, n (%)	12 (6.7)	13 (7.2)		
No formal diagnosis, but parent concerned, n (%)	38 (21.2)	52 (28.9)		
No diagnosis, n (%)	104 (58.1)	90 (50.0)		
Unanswered, n (%)	3 (1.7)	4 (2.2)		

[a] Two children who were reported by their parents to have a past diagnosis of autism or Asperger's syndrome were categorized under "Multiple diagnosis" as they also had another current mental health diagnosis.

Baseline Sample Characteristics, Attrition, and Symptom Elevation Follow-Up

Parent and child participants did not differ significantly between the two conditions on any of the outcome measures or demographic variables assessed at baseline. As shown in Table 2, most parent participants were female, married, or in a de facto relationship; living with the adolescent participant; and employed at least part-time. Almost 60% had attained graduate or postgraduate qualifications and reported either a past and/or current mental health diagnosis. Just over half of the adolescents were male, and most adolescents were reported by their parents to have no prior or current mental health diagnosis.

Adherence to the intervention and attrition rates for the follow-up assessment did not differ significantly between conditions and were not related to any participant characteristics at baseline except for parent age. Parents within dyads with missing postintervention data were younger than those with available postintervention data.

We also examined whether follow-up actions taken in response to elevations in adolescent depression and anxiety symptoms differed between conditions, as these may have impacted on reported outcomes at postintervention. Chi-square tests indicated no significant differences between conditions in the proportion of follow-up actions undertaken (*p* s>.05).

Time Interval Between Baseline and Postintervention Assessment Completion

The mean time interval between parent baseline and postintervention assessment completions was 118 days (SD=34.4, median=105, range=86-279). The mean time interval between child baseline and postintervention assessment completions was 110 days (SD=28.0, median=99, range=85-277 days). These time intervals did not differ significantly between conditions. The wide interval ranges were due to programming errors whereby automated postintervention assessment alerts for the research team to follow up with 59 dyads were not delivered. The error was detected during an audit of participant numbers on August 1, 2016, which also indicated that the follow-up assessment date for another dyad had been missed because of human error. Of the 60 dyads, 59 were contacted to complete the postintervention assessment; the remaining one was not contacted for this assessment as their final (12-month) follow-up assessment date was too near to this time (ie, within a month). The proportion of participants who were affected by this technical error did not differ between conditions.

Time Interval Between Parent and Child Assessment Completions at Each Assessment

For baseline assessments, the mean time interval between child and parent assessment completions, for dyads where both were completed (n=332) was 4.88 days (SD=8.62, range=0-76). For postintervention assessments, time intervals between completed child and parent assessments (n=294) averaged 11.4 days (SD=21.2, range=0-167). The extreme larger ends of these ranges were because of a small number of parent participants being difficult to contact during extended school holiday periods.

Intervention Completion and Adherence Rates Within the Whole Sample

At the time of data extraction, the average intended program usage within the intervention group (n=179) was 6.85 out of the nine modules available for selection in the personalized program. The average observed usage within the intervention group was 5.17 modules. Participants in the intervention group completed an average of 73.7% of their locked-in program. Intervention adherence within the intervention group was 44.1% (n=79). In addition, 15.1% of the intervention group (n=27) completed modules not initially selected as part of their personalized programs (average of 2 modules). Intervention adherence within the control group (n=180) was 72.8% (n=131).

Intervention Completion and Adherence Rates Within Follow-Up Sample

Of the 318 parents who completed the postintervention assessment, 165 (control=108, intervention=57) completed 100% of their program during the active intervention phase, which was defined as the time between completion of parent baseline assessment and completion of child postintervention assessment (a parent's follow-up completion timestamp was used if their child did not complete the postintervention assessment). Of the 308 child participants who completed the postintervention assessment, the parents of 159 (control=106, intervention=53) had completed their programs before completing their follow-up assessments.

A further 14 parent participants from the intervention group who completed their postintervention assessment did not have completion timestamps on some of their completed modules in their program, most likely because of technical error. Consequently, we were unable to conclusively determine whether they had completed their programs within the active intervention phase. There were 8 adolescent participants from this group who completed the postintervention assessment.

There were 23 parents (control=17, intervention=6) who completed their program after the active intervention phase. At the time of their follow-up assessment, intervention group parents in this subsample had completed an average of 53.0% of their program, whereas the average completion rate was 70.6% for control group parents.

The remaining 116 parent participants (control=39, intervention=77) who completed the postintervention assessment still had not completed their whole program at the time their data was extracted for analyses. In this subgroup, the average intervention completion rates were 61.9% and 68.7%, respectively, for intervention and control group parents.

Primary Outcome Analysis

Parenting to Reduce Adolescent Depression and Anxiety Scale (PRADAS), Parent-Report

We found a significant interaction between condition and time on total PRADAS scores, Cohen *d*=0.57, 95% CI 0.34-0.79. The intervention group's mean PRADAS score significantly exceeded that of the control group at postintervention, producing a small effect (see Table 3 and Figure 2; observed mean and SD are presented in Multimedia Appendix 4).

Secondary Outcome Analyses

Adolescent Anxiety and Depression

There were no significant interactions between condition and time on the SMFQ-P, SMFQ-C, and SCAS-P scores. Across conditions, parent participants reported significantly decreased symptoms of depression and anxiety from baseline to post intervention (Table 3).

A condition by time interaction on adolescent-reported anxiety was observed (Table 3; see also Figure 2), and individual comparisons revealed that self-reported anxiety decreased at postintervention only for the control group, with no change in the intervention group (Table 3). However, this interaction was no longer significant when the transformed SCAS-C data was analyzed; instead, both conditions showed significant decreases in symptoms of anxiety from baseline to postintervention (see Multimedia Appendix 4).

Parental Concordance With the Parenting Guidelines, Adolescent-Report

There was no significant condition by time interaction on the PRADAS-A total score. Instead, both groups reported significantly reduced parental concordance with the parenting guidelines from baseline to post intervention.

Post-Hoc Moderation Analyses

Parental Concordance With the Parenting Guidelines, Parent-Report

Post-hoc analyses revealed that adolescent-report depression and anxiety symptoms at baseline did not significantly moderate intervention effects on change in PRADAS scores from baseline to post intervention, $F_{1,294}$=3.59, *P*=.059, accounting for 1% variance. Conditional effects analyses revealed that among parents whose adolescents reported lower (SCAS-C≤12.42) and average (12.42<SCAS-C<46.68) levels of anxiety at baseline, PiP led to greater increases in PRADAS scores compared with the control condition (see Multimedia Appendix 4). The difference between conditions was not significant for parents whose adolescents reported higher levels of baseline anxiety (SCAS-C≤46.68).

Parental Concordance With the Parenting Guidelines, Adolescent-Report

Moderation analyses revealed that parent-reported adolescent anxiety moderated intervention effects on adolescent-report parenting, $F_{1,303}$=20.09, *P*<.001, accounting for 6.2% variance in change in PRADAS-A scores. Among parents who reported lower levels of adolescent anxiety at baseline (SCAS-P≤6.04), adolescents whose parents received PiP reported no change in PRADAS-A scores, whereas adolescents whose parents received the control intervention reported a reduction in PRADAS-A scores, *d*=0.50 (95% CI 0.27-0.72). However, among parents who reported higher levels of adolescent anxiety at baseline (SCAS-P≥29.30), the opposite pattern was observed, with PRADAS-A scores remaining stable in the control condition but decreasing in the PiP condition, *d*=−0.53 (95% CI −0.76 to −0.31]. No significant difference between conditions was observed at mean levels of parent-reported adolescent anxiety. Parent-reported adolescent depressive symptom levels at baseline did not moderate intervention effects on PRADAS-A scores. See Multimedia Appendix 4 for specific conditional effects referred to above.

Table 3. Estimates of marginal means (EMM), standard errors (SE), and mixed model repeated measures analyses of primary and secondary outcome scores at baseline and postintervention time points. There were no significant differences between conditions on any of the primary or secondary outcome measures at baseline.

Outcome measure and time	Outcome scores		*F* [a]	df	*P* value	d_{post} [b] (95% CI)	$d_{interaction}$ [c] (95% CI)
	Intervention EMM (SE)	Control EMM (SE)					
PRADAS[d]							
Baseline	46.40 (0.57)	47.40 (0.57)	25.54	1, 320	<.001	0.27 (0.05-0.49)	0.57 (0.34-0.79)
Postintervention	51.21 (0.62)	49.38 (0.61)					
PRADAS-A[e]							
Baseline	24.44 (0.44)	24.88 (0.43)	0.04	1, 308	.835	−0.11 (−0.34 to 0.11)	0.02 (−0.20 to 0.25)
Postintervention	23.40 (0.50)	23.94 (0.49)					
SCAS-P[f]							
Baseline	17.99 (0.90)	18.51 (0.89)	0.16	1, 327	.693	0.04 (−0.18 to 0.26)	0.04 (−0.18 to 0.26)
Postintervention	14.98 (0.89)	15.14 (0.88)					
SCAS-C[g]							
Baseline	28.73 (1.36)	30.20 (1.33)	5.08	1, 306	.025[h]	0.09 (−0.14 to 0.31)	0.26 (0.03-0.48)
Postintervention	28.17 (1.50)	26.75 (1.43)					
SMFQ-P[i]							
Baseline	5.07 (0.40)	4.75 (0.40)	0.33	1, 333	.566	0.05 (−0.17 to 0.27)	0.06 (−0.16 to 0.28)
Postintervention	3.51 (0.34)	3.47 (0.34)					
SMFQ-C[j]							
Baseline	6.16 (0.47)	6.40 (0.46)	0.26	1, 308	.609	0.04 (−0.19 to 0.26)	0.06 (−0.17 to 0.28)
Postintervention	6.16 (0.49)	6.14 (0.48)					

[a] Test of the condition by time interaction.

[b] Cohen *d* effect size of difference between conditions at postintervention, reported with 95% CI in parentheses.

[c] Cohen *d* effect size of interaction, reported with 95% CI in parentheses.

[d] PRADAS: Parenting to Reduce Adolescent Depression and Anxiety Scale.

[e] PRADAS-A: Parenting to Reduce Adolescent Depression and Anxiety Scale- Adolescent.

[f] SCAS-P: *Spence Children's Anxiety Scale*-Parent version.

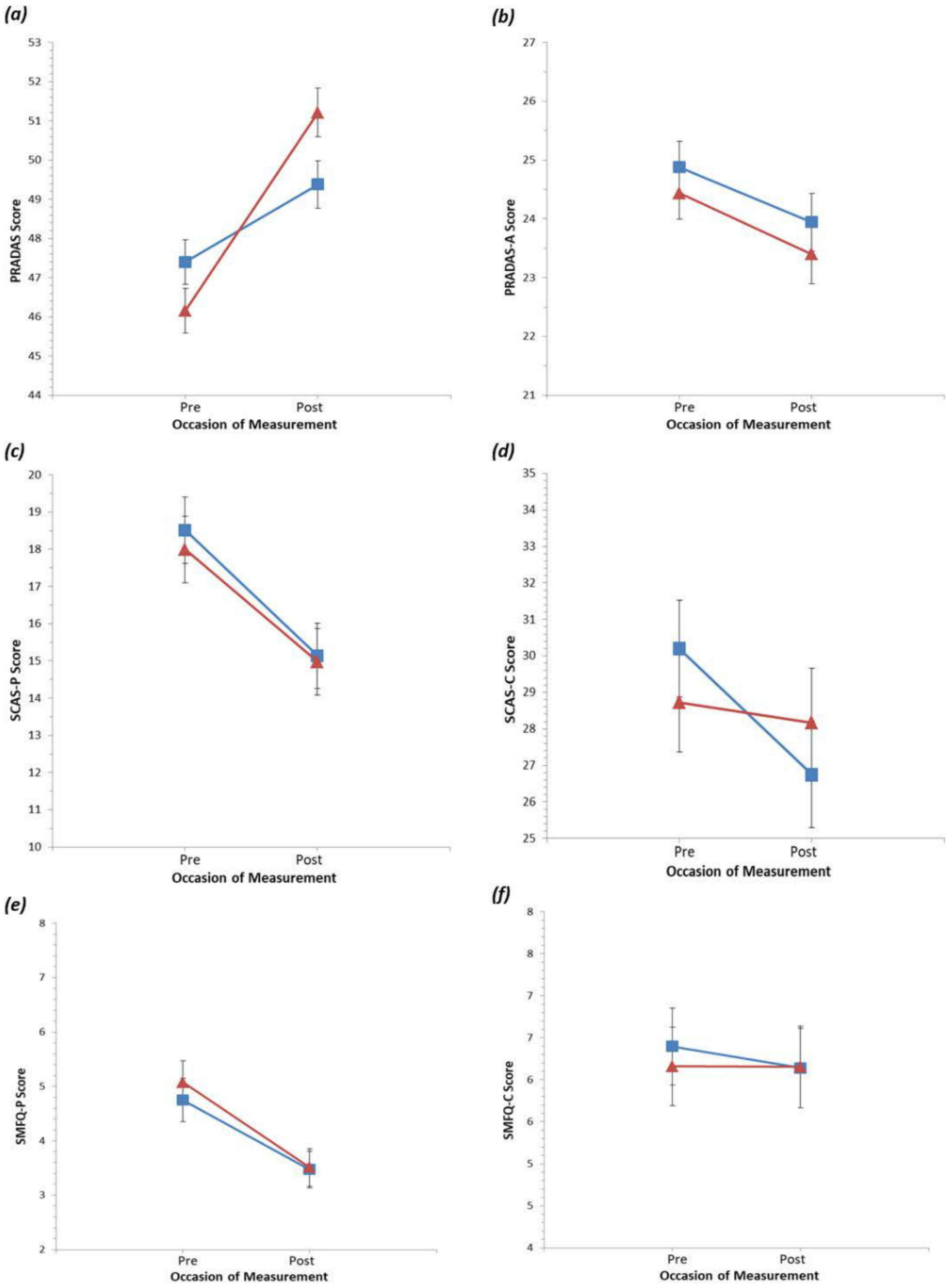
[g] SCAS-C: *Spence Children's Anxiety Scale*-Child version.

[h] This effect became nonsignificant when the analysis was run using log-transformed data to correct the skewed distribution.

[i] SMFQ-P: Short Moods and Feelings Questionnaire-Parent version.

[j] SMFQ-C: Short Moods and Feelings Questionnaire-Child version.

Figure 2. Graphs of estimated marginal means of primary and secondary outcomes for each intervention condition and occasion of measurement. Error bars represent standard error. Square=control, triangle=intervention.



Adolescent Anxiety and Depression

Adolescent-report baseline depressive symptoms had a significant moderation effect on change in SMFQ-P scores, $F_{1,294}$=5.97, P=.015, accounting for 1.9% variance. Conditional effects analyses revealed that change in SMFQ-P scores were significantly different between conditions only among adolescents who had reported higher levels of depressive symptoms at baseline (SMFQ-C≥12.19), with parents who received PiP reporting greater symptom reduction than those who received the control intervention, d=−0.35 (95% CI −0.58 to −0.12; see Multimedia Appendix 4). No significant moderation effects were found for changes in SMFQ-C, SCAS-C, and SCAS-P.

Discussion

Principal Findings: Primary Outcome

This RCT evaluated the short-term effects of the PiP intervention, a Web-based parenting intervention to prevent adolescent depression and anxiety. As hypothesized, compared with an educational-factsheet control intervention, PiP was found to produce greater improvements from baseline to post intervention (3 months later) in self-reported parenting behaviors, indicating increased concordance with evidence-based Parenting Guidelines [28]. This represented a medium effect size, although the intervention and control group difference at postintervention was a small effect. The parenting behaviors assessed represent parental risk and protective factors for adolescent depression and anxiety [12,29,34]; hence the current findings suggest that by modifying these parenting behaviors, PiP may have the potential to confer protection against adolescent depression and anxiety over the long term.

Principal Findings: Secondary Outcomes

Three months after commencing the intervention, no corresponding reductions in adolescent depressive and anxiety symptoms were found in secondary outcome analyses based on either parent or adolescent report. There was also no significant improvement in parenting based on adolescent report; in fact, across the sample, adolescents reported a slight reduction in guidelines-concordant parenting behaviors over the 3 months between the baseline and postintervention assessments. The lack of demonstrable effects of PiP on adolescent-reported parenting suggests that the parent-reported improvements in parenting were not noticed or not interpreted as such by their adolescents. This may reflect difficulties inherent in assessing early adolescents' perceptions of parenting, as they may rely on more generalized impressions of their parent's parenting and may not have the capacity to distinguish between specific, recent behaviors from general parenting practices over time [44]. Other possibilities are that adolescents may need more time before they notice improvements in their parents' parenting, or that parents' perceived changes in parenting did not translate into tangible behavioral changes. Parent-adolescent divergence in perceptions of parenting, especially in early adolescence, is normative and well-established in the developmental literature [45]. Developmental perspectives posit that adolescents' maturation processes, including autonomy seeking and individuation, give rise to a period of time when parents and adolescents experience the same interactions differently, which in turn account for the divergence in parent-adolescent perceptions of parenting [46]. Long-term follow-up and further research is required to elucidate the associations between these short-term perceptions of parenting and long-term adolescent depression and anxiety outcomes.

Post-Hoc Moderation Effects

As a universal prevention program involving a primarily community-based sample, we found that PiP did not demonstrate any short-term effect on adolescent depressive or anxiety symptoms, likely due to a "floor effect" (symptom levels were low at baseline for most participants). Interestingly, post-hoc moderation analyses suggest that PiP did have a significant effect on reducing parent-reported adolescent depressive symptoms for adolescents who had reported higher levels of symptoms at baseline. This suggests that PiP may be useful as an indicated intervention for parents of adolescents with elevated depressive symptoms. Compared with an active-control condition, PiP produced greater improvements in parent-reported parenting regardless of adolescent baseline depressive or anxiety symptoms. Post-hoc findings involving changes in adolescent-reported parenting are more challenging to interpret, suggesting that the extent to which adolescents perceived a reduction in guidelines-concordant parenting behaviors depended on the adolescents' parent-reported anxiety symptoms at baseline. Among adolescents with lower anxiety levels at baseline, those whose parents received PiP (as opposed to the control intervention) perceived less reduction in guidelines-concordant parenting. On the other hand, among adolescents with higher anxiety levels at baseline, those whose parents received PiP perceived greater reduction in guideline-concordant parenting. Generally, although these post-hoc findings are promising because of the exploratory nature of these analyses and the inconsistencies across measures, future research is required to test the efficacy of PiP as an indicated prevention intervention in a rigorously designed RCT.

Intervention Engagement

PiP was generally well received by parents, with a reasonably high mean proportion of intervention completion (74% of their personalized program, which comprised an average of 6.85 modules), with 15% of parents completing additional modules outside of their initially selected program. The current RCT did not include a comparison group which received the online PiP without phone support; hence, we cannot determine whether the phone calls aided in the intervention completion rates. However, the rates from a five-arm RCT of Web-based preventive interventions for adults [47] provide a basis for inferring that the phone calls contributed to higher rates of intervention completion. Specifically, that trial included three intervention arms (active website, active website with email, and active website with telephone) that achieved intervention completion rates of 37%, 55%, and 73%, respectively. Coincidentally, the proportion of PiP completion, which was supported by weekly administrative reminder phone calls, was very similar to the "active website with telephone" arm of the earlier RCT. If PiP was implemented without phone support, which conveys a sense of accountability [48], it is possible that completion rates may be closer to 40%. Nonetheless, qualitative feedback from parents, provided spontaneously through the weekly phone calls, indicate that most parents found PiP to be engaging and that the strategies provided were practical and useful. Some parents found it overwhelming to receive their personalized feedback report after completing their baseline assessment (about 45-min long) and then be required to start their first module immediately thereafter. In a real-world setting, this may not be such an issue as only the PRADAS, which takes less than 15 min to complete [34], would be required to generate the tailoring of PiP. Some parents, especially where there are existing challenges with the family situation (eg, heightened family conflict) or adolescent mental health, also sought more support and information through the phone calls. This suggests that if implemented as an indicated prevention program, phone support may need to be tailored to the parents' level of need.

Strengths and Limitations

This study had several notable strengths. It recruited a large community-based sample of parents and early adolescents and achieved a relatively low attrition rate for an online intervention at 3-month follow-up. Parents' engagement in their allocated intervention was acceptable in the intervention group and high in the control group. One methodological strength was the utilization of an active-control comparison group, which allowed us to test whether the PiP intervention yields benefits beyond simply involving parents in a parenting intervention [49]. Another strength was the inclusion of both parent and adolescent informants of parenting and adolescent symptoms, which permitted an exploration of outcome-reporting bias by parents who were aware of their group allocation; it also enabled analytic models (eg, in post-hoc analyses) that attempted to minimize shared method variance.

Nonetheless, there are limitations to this RCT. First, it recruited a large proportion of self-selected parents who were mostly mothers, highly educated, from intact families, and did not speak a language other than English at home. Although such sociodemographic characteristics are typical of adult participants in other online interventions [50] and are similar to a preventive online parenting intervention for preschool children [51], further research is required to determine whether PiP would have similar effects for families with different characteristics. Second, for reasons of parsimony, the current RCT only included one parent and one adolescent per family. In future research and outside the research context, it is possible for both parents to receive their own personalized intervention, which can help both of them become more concordant with the recommendations of the Parenting Guidelines and coparent more effectively. Existing evidence from programs that examined the unique effects of father and mother engagement in parenting interventions [52] suggests that including both parents could in turn enhance the effects of the intervention on child outcomes. Third, to reduce participant burden, we relied on new parenting measures developed specifically for this study, to the exclusion of other validated parenting measures, especially by adolescents who are subject to less outcome-reporting bias as they are less likely to be aware of their parent's intervention assignment. One implication of this is, for example, that we could not test whether the nonsignificant intervention effect on adolescent-report parenting is because of the insensitivity of the new PRADAS-A measure to detect changes in parenting, or broader developmental factors as discussed above [44-46]. We also did not collect observational data on parent-adolescent interactions, which would be useful for verifying self-reported changes in parenting. Although we used an active-control intervention, we did not assess or control for differential motivations and expectations in parents receiving each intervention (it is clear that most parents in the control condition wanted to receive PiP instead), hence, limiting our ability to draw firm causal conclusions [49]. Moreover, the control condition was not ideally matched in terms of interactivity and the amount of administrative support received (ie, on average, parents in this condition received fewer weekly phone calls than those in the intervention group). It was also not possible to determine whether parents in the control condition actually completed their intervention (ie, read all their factsheets). The "completion" of a factsheet may be better matched to the "completion" of a module if, for example, parents in the control condition were required to click on a button to "Finish session," just as parents in the intervention condition were required to click on a "Finish module" button. Future evaluations of PiP should also include other measures of adolescent functioning outcomes (eg, quality of life, emotion regulation, and school engagement) and more broadly, measures of cost-effectiveness.

Comparison With Prior Work

To the best of our knowledge, PiP is the first tailored Web-based parenting intervention for the prevention of adolescent depression and anxiety. A recent systematic review of RCTs of parenting interventions to prevent internalizing problems in children in the age range of 0 to 18 years revealed a dearth of rigorously evaluated, evidence-based interventions for parents of adolescents [9]. Of the 51 programs included in the review, only three were designed for parents of adolescents [53-55]. Moreover, despite the potential value of Web-based delivery for universal prevention programs [22,24], none of these three existing interventions are Web-based, and only one of these, the Tuning in to Teens (TINT; [54]) program, was found to successfully improve parenting and reduce adolescent anxiety (based on both parent and adolescent report) and depressive symptoms (based on parent report only) at approximately 9 months post intervention. Specifically, compared with a no-intervention control condition, TINT yielded a moderately large effect size on the change over time in parent-report parenting (Cohen d=0.76). Considering the differences in the control condition (no-intervention vs active-control) and modality (face-to-face vs Web-based), PiP's effects on parenting compare favorably with TINT's.

A limitation of existing preventive parenting programs is their focus on one or only a few parenting risk and protective factors for adolescent depression and anxiety [9]. As parents differ in their areas of competence and difficulties, such a narrow-focus approach may mean that these programs do not adequately address the range of modifiable risk and protective parenting factors for adolescent depression and anxiety that are relevant for each parent or family. As an automatically tailored program, PiP addresses this limitation by screening each parent across all evidence-based parenting risk and protective factors to provide a more thorough coverage of areas that may be important to target in preventive intervention for the particular parent.

Conclusions

The current RCT found that compared with an active-control condition, PiP produced greater short-term improvements in parental risk and protective factors for adolescent depression and anxiety from the parents' perspective, which represents the most proximal outcome. Moreover, the intervention was well received by parents. It remains to be ascertained, in longer-term follow-up (a 12-month follow-up assessment has commenced), whether the self-reported improvements in parenting will translate into corresponding adolescent-reported improvements and in turn give rise to protection against depression and anxiety in the adolescent. Given the brevity of the intervention and its fully developed Web-delivery modality, these preliminary findings are promising and suggest that PiP may be useful as a low-cost, scalable, and sustainable public health universal prevention program to empower parents in their parenting of adolescents.

Acknowledgments

The authors acknowledge funding from the National Health and Medical Research Council for the Web development of the PiP intervention, and the partnership of *beyondblue, the national depression and anxiety initiative* in the development of the parenting guidelines. The authors received salary support from the NHMRC for a Career Development Fellowship (MBHY, APP1061744) and a Senior Principal Research Fellowship (AFJ, APP1059785), an Australian Research Council Laureate Fellowship (RMR, FL150100096), and an Australian Government Research Training Program Scholarship (CN). The RCT is supported by an Australian Rotary Health Research Grant and Monash University's Advancing Women's Research Success Grant. None of the funding sources had any role in the conduct or publication of this study. The authors thank the reference group of parents and focus groups of students who contributed to the development of the PiP intervention. They also acknowledge with thanks the contributions of Jennifer Hanson-Peterson, Mairead Cardamone-Breen, and Jacqueline Green in the development and testing of the intervention and in project management, as well as their research assistants and the schools and organizations that assisted with recruitment for the RCT.

Conflicts of Interest

MBHY, AFJ, and KAL are codevelopers of the PiP intervention, and MBHY and AFJ are cofounders of the broader Parenting Strategies online platform of parenting interventions, including PiP. None of these authors derives a personal financial benefit from these online interventions.

Multimedia Appendix 1

CONSORT‐EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 678KB - jmir_v20i1e17_app1.pdf]

Multimedia Appendix 2

Participants' informed consent documentation.

[PDF File (Adobe PDF File), 698KB - jmir_v20i1e17_app2.pdf]

Multimedia Appendix 3

Screenshots of the Partners in Parenting intervention.

[PDF File (Adobe PDF File), 856KB - jmir_v20i1e17_app3.pdf]

Multimedia Appendix 4

Observed means and standard deviations of outcome variables, results of mixed effect model repeat measurements (MMRMs) using transformed data, and posthoc analyses.

[PDF File (Adobe PDF File), 66KB - jmir_v20i1e17_app4.pdf]

References

1. Kessler RC, Petukhova M, Sampson NA, Zaslavsky AM, Wittchen H. Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. Int J Methods Psychiatr Res 2012 Sep;21(3):169-184 [FREE Full text] [doi: 10.1002/mpr.1359] [Medline: 22865617]
2. Lawrence D, Johnson S, Hafekost J, Boterhoven DHK, Sawyer M, Ainley J. Department of Health. Canberra; 2015. The Mental Health of Children and Adolescents: Report on the second Australian Child and Adolescent Survey of Mental Health and Wellbeing URL: https://www.health.gov.au/internet/main/publishing.nsf/Content/9DA8CA21306FE6EDCA257E2700016945/%24File/child2.pdf [accessed 2017-12-30] [WebCite Cache ID 6w6xWzkLU]
3. Rao U, Ryan ND, Birmaher B, Dahl RE, Williamson DE, Kaufman J, et al. Unipolar depression in adolescents: clinical outcome in adulthood. J Am Acad Child Adolesc Psychiatry 1995 May;34(5):566-578. [doi: 10.1097/00004583-199505000-00009] [Medline: 7775352]

4. Last CG, Hansen C, Franco N. Anxious children in adulthood: a prospective study of adjustment. J Am Acad Child Adolesc Psychiatry 1997 May;36(5):645-652. [doi: 10.1097/00004583-199705000-00015] [Medline: 9136499]
5. Andrews G, Issakidis C, Sanderson K, Corry J, Lapsley H. Utilising survey data to inform public policy: comparison of the cost-effectiveness of treatment of ten mental disorders. Br J Psychiatry 2004 Jun 01;184(6):526-533. [doi: 10.1192/bjp.184.6.526]
6. Patel V, Flisher A, Hetrick S, McGorry P. Mental health of young people: a global public-health challenge. Lancet 2007 Apr;369(9569):1302-1313. [doi: 10.1016/S0140-6736(07)60368-7]
7. Fisak BJ, Richard D, Mann A. The prevention of child and adolescent anxiety: a meta-analytic review. Prev Sci 2011 Sep;12(3):255-268. [doi: 10.1007/s11121-011-0210-0] [Medline: 21437675]
8. Merry SN, Hetrick SE, Cox GR, Brudevold-Iversen T, Bir JJ, McDowell H. Psychological and educational interventions for preventing depression in children and adolescents. Cochrane Database Syst Rev 2011 Dec 07(12):CD003380. [doi: 10.1002/14651858.CD003380.pub3] [Medline: 22161377]
9. Yap MB, Morgan AJ, Cairns K, Jorm AF, Hetrick SE, Merry S. Parents in prevention: a meta-analysis of randomized controlled trials of parenting interventions to prevent internalizing problems in children from birth to age 18. Clin Psychol Rev 2016 Oct 21;50:138-158. [doi: 10.1016/j.cpr.2016.10.003] [Medline: 27969003]
10. Beesdo K, Knappe S, Pine DS. Anxiety and anxiety disorders in children and adolescents: developmental issues and implications for DSM-V. Psychiatr Clin North Am 2009 Sep;32(3):483-524 [FREE Full text] [doi: 10.1016/j.psc.2009.06.002] [Medline: 19716988]
11. Cairns KE, Yap MB, Pilkington PD, Jorm AF. Risk and protective factors for depression that adolescents can modify: a systematic review and meta-analysis of longitudinal studies. J Affect Disord 2014 Dec;169:61-75. [doi: 10.1016/j.jad.2014.08.006] [Medline: 25154536]
12. 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 Jan;62:299-329 [FREE Full text] [doi: 10.1146/annurev.psych.121208.131619] [Medline: 20822438]
13. Yap MB, Pilkington PD, Ryan SM, Jorm AF. Parental factors associated with depression and anxiety in young people: a systematic review and meta-analysis. J Affect Disord 2014 Mar;156:8-23. [doi: 10.1016/j.jad.2013.11.007] [Medline: 24308895]
14. Yap MB, Jorm AF. Parents' beliefs about actions they can take to prevent depressive disorders in young people: results from an Australian national survey. Epidemiol Psychiatr Sci 2011 Sep 22;21(01):117-123. [doi: 10.1017/S2045796011000667]
15. Fisak BJ, Richard D, Mann A. The prevention of child and adolescent anxiety: a meta-analytic review. Prev Sci 2011 Sep;12(3):255-268. [doi: 10.1007/s11121-011-0210-0] [Medline: 21437675]
16. Hetrick SE, Cox GR, Witt KG, Bir JJ, Merry SN. Cognitive behavioural therapy (CBT), third-wave CBT and interpersonal therapy (IPT) based interventions for preventing depression in children and adolescents. Cochrane Database Syst Rev 2016 Aug 09(8):CD003380. [doi: 10.1002/14651858.CD003380.pub4] [Medline: 27501438]
17. 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]
18. Mrazek P, Haggerty R. Reducing Risks for Mental Disorders: Frontiers for Preventive Intervention Research. Washington DC: National Academy Press; 1994.
19. Rose G. The Strategy of Preventive Medicine. Oxford: Oxford University Press; 1992.
20. Koerting J, Smith E, Knowles MM, Latter S, Elsey H, McCann DC, et al. Barriers to, and facilitators of, parenting programmes for childhood behaviour problems: a qualitative synthesis of studies of parents' and professionals' perceptions. Eur Child Adolesc Psychiatry 2013 Nov;22(11):653-670 [FREE Full text] [doi: 10.1007/s00787-013-0401-2] [Medline: 23564207]
21. Australian Bureau of Statistics. Canberra, Australia: Australian Bureau of Statistics; 2014. 8146.0 - Household Use of Information Technology URL: http://www.abs.gov.au/ausstats/abs@.nsf/PrimaryMainFeatures/8146.0?OpenDocument [accessed 2017-05-22] [WebCite Cache ID 6qdiS9O1z]
22. Cuijpers P, van Straten A, Warmerdam L, van Rooy M. Recruiting participants for interventions to prevent the onset of depressive disorders: possible ways to increase participation rates. BMC Health Serv Res 2010;10:181 [FREE Full text] [doi: 10.1186/1472-6963-10-181] [Medline: 20579332]
23. 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]
24. Yap MB, Martin PD, Jorm AF. Online parenting guidelines to prevent adolescent depression and anxiety: evaluating user characteristics and usefulness. Early Interv Psychiatry 2017 Oct 27:- Epub ahead of print(forthcoming). [doi: 10.1111/eip.12503] [Medline: 29076246]
25. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. PLoS One 2010;5(10):e13196 [FREE Full text] [doi: 10.1371/journal.pone.0013196] [Medline: 20967242]

26. Love SM, Sanders MR, Turner KM, Maurange M, Knott T, Prinz R, et al. Social media and gamification: engaging vulnerable parents in an online evidence-based parenting program. Child Abuse Negl 2016 Mar;53:95-107. [doi: 10.1016/j.chiabu.2015.10.031] [Medline: 26880281]
27. Yap MB, Lawrence KA, Rapee RM, Cardamone-Breen MC, Green JM, Jorm AF. Partners in parenting: a multi-level web-based approach to support parents in prevention and early intervention for adolescent depression and anxiety. JMIR Ment Health 2017 Dec 19;4(4):e59. [doi: 10.2196/mental.8492] [Medline: 29258974]
28. Parenting Strategies Program. Melbourne: BeyondBlue; 2013. How to prevent depression and clinical anxiety in your teenager: Strategies for parents URL: https://www.parentingstrategies.net/depression/media/pdf/Guidelines.pdf [accessed 2017-12-30] [WebCite Cache ID 6w6v5ZOmi]
29. Jorm AF. Using the Delphi expert consensus method in mental health research. Aust N Z J Psychiatry 2015 Aug 21;49(10):887-897. [doi: 10.1177/0004867415600891]
30. Sanders MR, Kirby JN. Consumer engagement and the development, evaluation, and dissemination of evidence-based parenting programs. Behav Ther 2012 Jun;43(2):236-250 [FREE Full text] [doi: 10.1016/j.beth.2011.01.005] [Medline: 22440062]
31. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. CAIS 2009;24:485-500 [FREE Full text]
32. 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 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
33. Kreuter M, Farrell D, Olevitch L, Brennan L. Tailoring Health Messages: Customizing Communication With Computer Technology. New Jersey: Lawrence Erlbaum Associates; 2000.
34. Cardamone-Breen MC, Jorm AF, Lawrence KA, Mackinnon AJ, Yap MB. The parenting to reduce adolescent depression and anxiety scale: assessing parental concordance with parenting guidelines for the prevention of adolescent depression and anxiety disorders. PeerJ 2017;5:e3825 [FREE Full text] [doi: 10.7717/peerj.3825] [Medline: 28951815]
35. Parenting Strategies Program. 2010. Parenting Strategies: Protecting Your Child's Mental Health URL: http://parentingstrategies.net/(Archived by WebCite® at http://www.webcitation.org/6rYJbsD6f) [accessed 2017-06-28] [WebCite Cache ID 6rYJbsD6f]
36. Raising Children Network. Raising Children Network website - Teens URL: http://raisingchildren.net.au/teens/teens.html (Archived by WebCite® at http://www.webcitation.org/6qz1JBTsj) [accessed 2017-06-05] [WebCite Cache ID 6qz1JBTsj]
37. Angold A, Costello E, Messer S. Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. Int J Methods Psychiatr Res 1995;5:237-249.
38. Spence SH. Structure of anxiety symptoms among children: a confirmatory factor-analytic study. J Abnorm Psychol 1997;106(2):280-297. [doi: 10.1037/0021-843X.106.2.280]
39. Nauta MH, Scholing A, Rapee RM, Abbott M, Spence SH, Waters A. A parent-report measure of children’s anxiety: psychometric properties and comparison with child-report in a clinic and normal sample. Behav Res Ther 2004 Jul;42(7):813-839. [doi: 10.1016/S0005-7967(03)00200-6]
40. 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 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
41. Gueorguieva R, Krystal JH. Move over ANOVA: progress in analyzing repeated-measures data and its reflection in papers published in the Archives of General Psychiatry. Arch Gen Psychiatry 2004 Mar;61(3):310-317. [doi: 10.1001/archpsyc.61.3.310] [Medline: 14993119]
42. Hayes AF. AFHayes.com. 2012. PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling URL: http://afhayes.com/public/process2012.pdf [accessed 2017-12-30] [WebCite Cache ID 6w6wxvymt]
43. Bell ML, Fairclough DL. Practical and statistical issues in missing data for longitudinal patient-reported outcomes. Stat Methods Med Res 2014 Oct;23(5):440-459. [doi: 10.1177/0962280213476378] [Medline: 23427225]
44. Branje SJ, van Aken MA, van Lieshout CF. Relational support in families with adolescents. J Fam Psychol 2002;16(3):351-362. [doi: 10.1037/0893-3200.16.3.351]
45. Branje S, Laursen B, Collins W. Parent-child communication during adolescence. In: Routledge handbook of family communication. 2nd edition. New York: Routledge; 2012:271-286.
46. Collins W, Laursen B. Adolescent relationships: The art of fugue. In: Close Relationships: A Sourcebook. Thousand Oaks, CA: Sage Publications; 2000:59-69.
47. Christensen H, Batterham P, Mackinnon A, Griffiths KM, Kalia HK, Kenardy J, et al. Prevention of generalized anxiety disorder using a web intervention, iChill: randomized controlled trial. J Med Internet Res 2014;16(9):e199 [FREE Full text] [doi: 10.2196/jmir.3507] [Medline: 25270886]
48. 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 Mar;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]

49. Boot WR, Simons DJ, Stothart C, Stutts C. The pervasive problem with placebos in psychology: why active control groups are not sufficient to rule out placebo effects. Perspect Psychol Sci 2013 Jul;8(4):445-454. [doi: 10.1177/1745691613491271] [Medline: 26173122]
50. Murray E. Web-based interventions for behavior change and self-management: potential, pitfalls, and progress. Med 2 0 2012 Aug;1(2):e3 [FREE Full text] [doi: 10.2196/med20.1741] [Medline: 25075231]
51. Morgan AJ, Rapee RM, Salim A, Goharpey N, Tamir E, McLellan LF, et al. Internet-delivered parenting program for prevention and early intervention of anxiety problems in young children: randomized controlled trial. J Am Acad Child Adolesc Psychiatry 2017 May;56(5):417-425.e1. [doi: 10.1016/j.jaac.2017.02.010] [Medline: 28433091]
52. Frank TJ, Keown LJ, Sanders MR. Enhancing father engagement and interparental teamwork in an evidence-based parenting intervention: a randomized-controlled trial of outcomes and processes. Behav Ther 2015 Nov;46(6):749-763. [doi: 10.1016/j.beth.2015.05.008] [Medline: 26520218]
53. Bearslee WR, Wright EJ, Gladstone TR, Forbes P. Long-term effects from a randomized trial of two public health preventive interventions for parental depression. J Fam Psychol 2007 Dec;21(4):703-713. [doi: 10.1037/0893-3200.21.4.703] [Medline: 18179342]
54. Kehoe CE, Havighurst SS, Harley AE. Tuning in to teens: improving parent emotion socialization to reduce youth internalizing difficulties. Soc Dev 2013 Oct 15;23(2):413-431. [doi: 10.1111/sode.12060]
55. Rotheram-Borus MJ, Stein JA, Lester P. Adolescent adjustment over six years in HIV-affected families. J Adolesc Health 2006 Aug;39(2):174-182. [doi: 10.1016/j.jadohealth.2006.02.014] [Medline: 16857528]

Abbreviations

EMM: estimates of marginal means
MMRM: mixed model of repeated measures
PiP: Partners in Parenting
PRADAS: Parenting to Reduce Adolescent Depression and Anxiety Scale
PRADAS-A: Parenting to Reduce Adolescent Depression and Anxiety Scale-Adolescent version
PSD: Persuasive Systems Design
RCT: randomized controlled trial
SCAS: Spence Children's Anxiety Scale
SCAS-C: Spence Children's Anxiety Scale-Child version
SCAS-P: Spence Children's Anxiety Scale-Parent version
SMFQ: Short Mood and Feelings Questionnaire
SMFQ-C: Short Mood and Feelings Questionnaire-Child version
SMFQ-P: Short Mood and Feelings Questionnaire-Parent version
TINT: Tuning in to Teens

Edited by G Eysenbach; submitted 09.10.17; peer-reviewed by A Gulliver, Y Perry; comments to author 02.11.17; revised version received 29.11.17; accepted 05.12.17; published 19.01.18.

Please cite as:
Yap MBH, Mahtani S, Rapee RM, Nicolas C, Lawrence KA, Mackinnon A, Jorm AF
A Tailored Web-Based Intervention to Improve Parenting Risk and Protective Factors for Adolescent Depression and Anxiety Problems: Postintervention Findings From a Randomized Controlled Trial
J Med Internet Res 2018;20(1):e17
URL: http://www.jmir.org/2018/1/e17/
doi:10.2196/jmir.9139
PMID:29351895

Original Paper

Internet-Based Cognitive Behavioral Therapy for Children and Adolescents With Dental Anxiety: Open Trial

Shervin Shahnavaz[1], PhD; Erik Hedman-Lagerlöf[2], PhD; Tove Hasselblad[1], MSc (Psychology); Lena Reuterskiöld[3], PhD; Viktor Kaldo[4], PhD; Göran Dahllöf[1], PhD

[1]Division of Pediatric Dentistry, Department of Dental Medicine, Karolinska Institutet, Huddinge, Sweden
[2]Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
[3]Department of Psychology, Stockholm University, Stockholm, Sweden
[4]Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

Corresponding Author:
Shervin Shahnavaz, PhD
Division of Pediatric Dentistry
Department of Dental Medicine
Karolinska Institutet
Box 4064
14104
Huddinge,
Sweden
Phone: 46 8 52488091
Email: shervin.shahnavaz@ki.se

Abstract

Background: Cognitive behavioral therapy (CBT) is an evidence-based method for treating specific phobias, but access to treatment is difficult, especially for children and adolescents with dental anxiety. Psychologist-guided Internet-based CBT (ICBT) may be an effective way of increasing accessibility while maintaining treatment effects.

Objective: The aim of this study was to test the hypothesis that psychologist-guided ICBT improves school-aged children's and adolescents' ability to manage dental anxiety by (1) decreasing avoidance and affecting the phobia diagnosis and (2) decreasing the dental fear and increasing the target groups' self-efficacy. The study also aimed to examine the feasibility and acceptability of this novel treatment.

Methods: This was an open, uncontrolled trial with assessments at baseline, posttreatment, and the 1-year follow-up. The study enrolled and treated 18 participants. The primary outcome was level of avoidance behaviors, as measured by the picture-guided behavioral avoidance test (PG-BAT). The secondary outcome was a diagnostic evaluation with the parents conducted by a psychologist. The specific phobia section of the structured interview Kiddie-Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime (K-SADS-PL) was used. Other outcome measures included level of dental anxiety and self-efficacy. The ICBT, which employed exposure therapy, comprised 12 modules of texts, animations, dentistry-related video clips, and an exercise package (including dental instruments). Participants accessed the treatment through an Internet-based treatment platform and received Web-based guidance from a psychologist. Treatment also included training at dental clinics. Feasibility and acceptability were assessed by measures of engagement, adherence, compliance, completed measures, patient and parent satisfaction scale, and staff acceptability.

Results: The level of avoidance (according to the primary outcome measure PG-BAT) and dental anxiety decreased and self-efficacy increased significantly ($P<.001$), within-group effect sizes for both the primary outcome (Cohen d=1.5), and other outcomes were large in the range of 0.9 and 1.5. According to K-SADS-PL, 53% (8/15) of the participants were free from diagnosable dental anxiety at the 1-year follow-up. At the 1-year follow-up, improvements were maintained and clinically significant, with 60% (9/15) of participants who had been unable to manage intraoral injection of local anesthetics before ICBT reporting having accomplished this task at a dental clinic. The target group showed improvement in all the outcome measures. High levels of feasibility and acceptability were observed for the treatment.

Conclusions: ICBT is a promising and feasible treatment for dental anxiety in children and adolescents. Integrating it into routine pediatric dental care would increase access to an effective psychological treatment. The results of this open trial must be replicated in controlled studies.

(J Med Internet Res 2018;20(1):e12) doi:10.2196/jmir.7803

KEYWORDS

cognitive behavioral therapy; dental fear; dental phobia; dentistry; internet-based treatment; pediatric dentistry; psychology; self efficacy

Introduction

Background

Dental fear and anxiety is defined as strong negative feelings associated with dental treatment or anticipation of dental treatment. Among children and adolescents, the prevalence of dental fear and anxiety is approximately 9% [1]. The Diagnostic and Statistical Manual of Mental Disorders-4th edition (DSM-IV), often used in psychiatric or psychological research, classifies dental anxiety as a form of specific phobia and defines it as a persistent, irrational, and intense fear of a specific object or medical procedure in dentistry persisting for at least 6 months [2]. Dental anxiety often begins during childhood or adolescence. It leads to poor oral health manifesting as untreated caries, missing teeth, or periodontal problems [3] and can have other negative consequences such as a sense of embarrassment and reduced self-esteem [4].

The common methods for dealing with dental anxiety in pediatric dentistry are tell-show-do, sedation with midazolam, nitrous oxide sedation, and general anesthesia [5,6]. According to a recent systematic review of methods in pediatric dentistry [7], the evidence supporting these methods is low, and it is uncertain whether they reduce dental anxiety. This highlights the need for new, evidence-based psychological methods for treating dental anxiety in pediatric dentistry [8,9].

Cognitive Behavioral Therapy

Cognitive behavioral therapy (CBT) is an evidence-based treatment method for psychiatric conditions such as specific phobias [10,11]. Main features of CBT include psychoeducation, coping techniques, cognitive restructuring, exposure, and homework exercises. CBT has been shown to be highly effective in adults with dental anxiety [12,13]. In a recently conducted randomized clinical trial, our research group showed that CBT has similarly large effects in children and adolescents [14]. Results from a qualitative study show that children experience increased feelings of safety and mastery and reduced fear in dental situations after receiving CBT [9,15]. However, accessibility to treatment is low. Children and parents face challenges that make it difficult to receive face-to-face CBT, such as constraints in time and availability, long distances to specialist pediatric dental clinics or lack of a psychologist at dental clinics. Thus, children with dental anxiety need better access to evidence-based psychological treatments.

Internet-Based Cognitive Behavioral Therapy

Internet-based cognitive behavioral therapy (ICBT) is based on the same principles as conventional CBT, although ICBT is delivered over the Internet instead of in face-to-face sessions. ICBT has effect sizes comparable with face-to-face CBT and has been scientifically evaluated for many psychiatric conditions in both adults and children [16,17]. ICBT has shown promising results in treating specific phobias in children [18-20], is easier to deliver, and is more cost-effective than face-to-face CBT. The Internet-based version of CBT thus improves access to CBT among children and young individuals. An open study of self-help CBT resources (available in paper-based and on-line versions) for children with dental anxiety showed that CBT resources are a feasible and acceptable intervention for the reduction of dental anxiety in children in the age range of 9 to 16 years [21]. The self-help CBT resources mentioned above lack therapist guidance. The ICBT with therapist guidance (psychologist contact on weekly basis through a chat system) we implemented in this study is an Internet adapted version of a previously evaluated face-to-face CBT program (including therapy sessions with a clinical psychologist on weekly basis). The face-to-face treatment program has been evaluated in a randomized controlled trial (RCT) study [14]. We produced the treatment manual for the ICBT program based on the face-to-face CBT manual, which we developed further and adapted for the Internet.

The aim of this study was to test the hypothesis that psychologist-guided ICBT improves school-aged children's and adolescents' ability to manage dental anxiety by (1) decreasing avoidance and affecting the phobia diagnosis and (2) decreasing the dental fear and increasing the target groups' self-efficacy. The study also aimed to examine the feasibility and acceptability of this novel treatment.

Methods

Design

This study has a single-group, open-trial design. We conducted assessments at baseline, posttreatment, and the 1-year follow-up. All participants and parents (one parent if there was only one primary caregiver) provided written informed consent.

To be able to have a larger recruitment base, we chose to include both children and teenagers in the study. As children needed to be able to read and understand the written text on the Internet platform, we chose 8 years as the minimum age. Age 7 or 8 is often the set starting age for these types of interventions. Some studies also have similar age ranges (7-13 or 7-14 years) [22,23]. The maximum age was chosen based on the need for self-determination and integrity that teenagers from the age of 15 years are granted in Sweden. If we would have included patients older than 15 years, then we would have to create two different log-ins in the Internet platform for the parent and the adolescents, which would make the treatment administration more complicated and would reduce the feasibility of the treatment. In the current intervention, participants and their parents had a common log-in and access to all modules throughout the course of treatment. The material on the Internet platform was common to all ages participating in the study, but the psychologist guiding the child tailored his or her messages to the child's age.

The regional ethics review board in Stockholm approved the study (Daybook no: 2014/633-31/5).

Participants and Recruitment

We recruited participants in two phases from August 2014 to February 2015. Our team contacted both private (only in Stockholm) and public dental clinics in Sweden by email and encouraged them to advertise the study in their waiting rooms (posters could be ordered or downloaded from the website). Interested parents applied through the website of the Department of Dental Medicine at Karolinska Institutet (the website provided visitors with brief information about the study and a list of primary inclusion criteria (items 1-5, Textbox 1). Textboxes 1 and 2 list the complete eligibility and exclusion criteria for participation in the study. A history of unsuccessful CBT for dental anxiety (during the past 3 years) suggests that this treatment might not be suitable. Therefore, we found it unethical to offer a treatment based on previous nonimprovement and excluded participants with earlier CBT experience. This is a standardized procedure in psychological treatment studies [24]

In all, 34 parents applied to participate in the study, and 18 children fulfilled all inclusion criteria and no exclusion criteria (we included a patient with attention deficit hyperactivity disorder incorrectly in the study; this patient received treatment and was kept in the analysis on the basis of intention-to-treat principle). Figure 1 shows the flow of participants through the trial.

The duration of dental fear before inclusion in the study was determined by directing the question "How many months has your child had his or her dental fear (including intraoral injection phobia)?" to the parent during the online screening (children and adolescents also received the question). The intensity of dental phobia was determined by 0 to 10 on a visual analog scale (VAS; no fear-strong fear) and the fear scale; Children's Fear Survey Schedule-Dental Subscale (CFSS-DS). Intensity equal to or less than 3 on VAS and values less or equal to 31 on CFSS-DS evaluated by both child and parent were considered as too weak fear, which led to exclusion. Evaluation of duration and intensity of dental anxiety were also part of the diagnostic instrument K-SADS-PL, which were used during the telephone interview with the parents, evaluated by a psychologist.

Textbox 1. Inclusion criteria for participation in this study.

1. The participant is in the age range of 8 to 15 years.
2. The participant had strong dental fear for at least 6 months before registering for the study.
3. The participant and parent have regular access to a computer and the Internet.
4. The participant is able to read and write in Swedish.
5. The participant has no current or planned psychological examination or treatment.
6. Participant and parents agree to participate in the research project.
7. The participant and parent have the time, opportunity, and motivation to work on and practice ICBT for 3 hours a week over 12 weeks.
8. Parents agree to book at least four visits at the dentist's office during the 12 weeks of treatment.
9. Parents give their consent for the participant to be exposed to intraoral injection at the dentist if the child suffers from intraoral injection phobia, even if the child has no dental treatment needs.

Textbox 2. Exclusion criteria for participation in this study.

1. Full scores on both the child and parent versions of the picture-guided behavioral avoidance test, which means that the child manages most of the procedures in dentistry.
2. A score of 31 or less on both the child and parent versions of the Children's Fear Survey Schedule- Dental Subscale while, at the same time, not fulfilling the criteria for intraoral injection phobia.
3. Likely fulfillment of the criteria for a neurodevelopmental disorder according to the Development and Well-being Assessment and/or telephone interview by a psychologist.
4. Other psychiatric disorders such as severe depression, an eating disorder, or self-harm behavior that need treatment before dentistry-related specific phobia.
5. Stressful life experiences during the past 12 months, such as a difficult divorce in the family or somatic illness that the parent or psychologist sees as an obstacle to treatment.
6. A history of cognitive behavioral treatment for dental anxiety or needle phobia during the past 3 years.

Figure 1. Flow of participants through the trial.

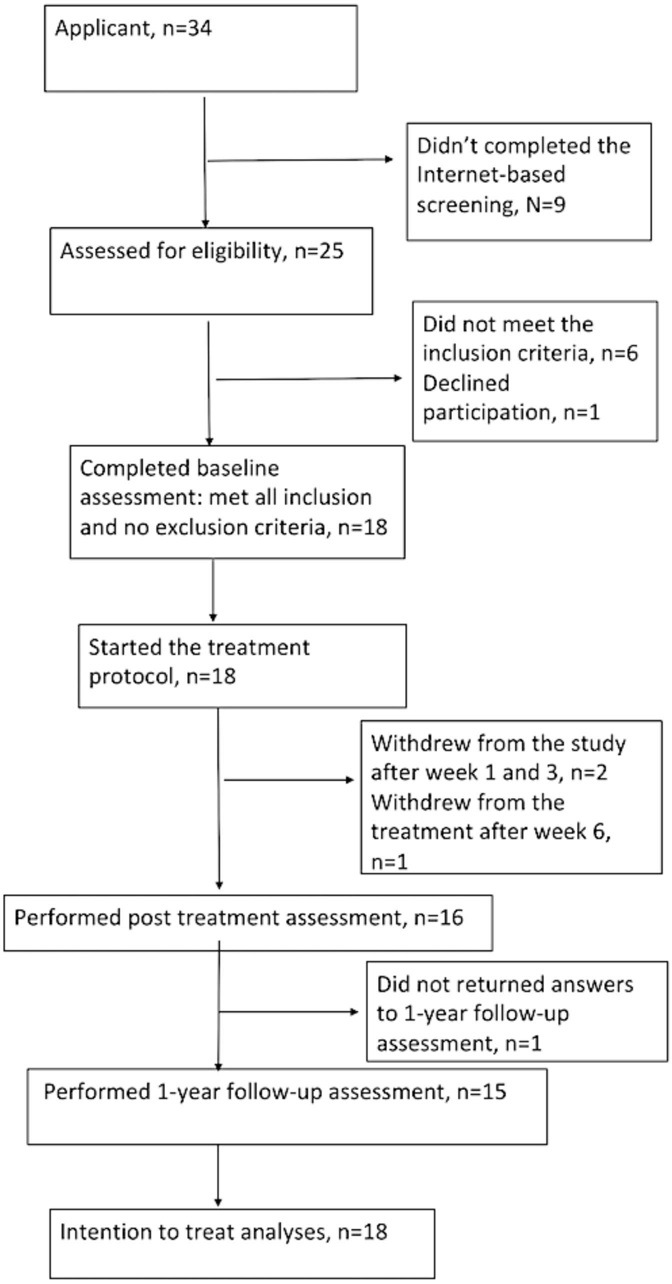


Table 1 presents the sociodemographic and clinical characteristics of the participants in the intervention group.

Primary Outcome Measure

The primary outcome measures were the child and parental versions of the picture-guided behavior avoidance test (PG-BAT). The PG-BAT is an analogous test to the behavior avoidance test (BAT), which is the recommended outcome measure for CBT studies of phobias [13,25,26]. The test is constructed by our research group and has been used in an earlier published study of face-to-face CBT for children and adolescents with dental phobia [14].

Our study chose the PG-BAT because it is adapted to pediatric dentistry and is self-reported, possible to conduct on Internet, in contrast to the BAT, which is administered by clinicians face-to-face. The PG-BAT consists of 17 hierarchically organized dental-clinical situations. During this online self-reported test, the participant views pictures of a child going through different dental procedures. The participant responds, using yes or no answers, whether he or she could manage the situation. The test comprises these situations: going into the dental treatment room, sitting in the treatment chair, fastening a paper bib around the neck, lowering back the chair, opening mouth, undergoing a clinical exam with mirror and dental probe, dropbox receiving topical anesthesia, receiving an injection of local anesthesia, undergoing drilling, and undergoing extraction. The same process applies for the parental version (parents assess the child's ability to manage the dental procedures listed above).

Table 1. Patient and clinical characteristics (n=18).

Characteristic	Value
Patient characteristics	
Age (years), mean (SD)	11 (2)
Gender (female), n (%)	11 (61)
Parent or sibling with dental fear, n (%)	3 (6)
University education, mother, n (%)	7 (39)
University education, father, n (%)	8 (47)
One or both parents born in a country other than Sweden, n (%)	3 (17)
Parents live together, n (%)	12 (67)
Clinical characteristics	
Comorbidity[a], n (%)	6 (30)
Duration of dental anxiety (years), mean (SD)	4 (3)
Intraoral injection as main fear, n (%)	14 (78)
Main reason for onset of fear	
Negative experiences in dentistry, n (%)	3 (17)
Negative experiences in health care, n (%)	11 (61)
Model learning through parent or sibling, n (%)	2 (11)
Do not know the reason, n (%)	2 (11)
Experience of nitrous oxide, midazolam, or general anesthesia before ICBT[b], n (%)	11 (60)
Referred to pediatric dental clinic, n (%)	8 (44)

[a] Comorbidity diagnoses were phobia of dogs, wasps, or blood, as well as attention deficit hyperactivity disorder and language disorder.

[b] ICBT: Internet-based cognitive behavioral therapy.

The score for each child was the sum of positive responses to these items, from 0 (not entering the dentist's room) to 17 (managing all activities including dental extraction), with 1 point for each stage until the point where the child discontinued the test. According to data (N=26) generated in an ongoing study, both the child and parental versions of this test have shown good internal reliability (with a Cronbach alpha of. 88 and .86, respectively) and validity (significant positive association between the instruments and a clinician-conducted face-to-face BAT; r=.68, N=36; P<.001 for the child version and r=.75, N=37; P<.001 for the parental version).

Secondary Outcome Measure

Diagnosis of dental phobia assessed the presence of dentistry-related specific phobia as measured by the phobic disorders supplement included in K-SADS-PL. K-SADS-PL is a semistructured diagnostic interview guide that generates reliable and valid psychiatric diagnoses [27]. We conducted interviews by telephone with one parent of the participants. All other outcome measures were self- or parent-reported and administered online.

Other Outcome Measures

CFSS-DS child and parental versions, which consist of 15 items measuring the degree of fear associated with various situations in dental and medical care and with interactions with people unfamiliar to the child (scale 1-5, from no fear to high fear). The CFSS-DS for the Swedish version of the test has high test-retest reliability and validity [28].

Self-Efficacy Questionnaire for Specific Phobias (SEQ-SP), which consists of 14 questions assessing the level of self-efficacy (a 5-point scale, with the endpoints 1=low self-efficacy and 5=high self-efficacy). We used a version we translated to Swedish and adapted to dentistry. Flatt and King [29] provide evidence for the reliability and validity of the SEQ-SP.

Children's negative cognitions in dentistry (CNCD), as measured using a scale constructed by our research team and inspired by a scale used in an earlier study of CBT for adults in dentistry [13]. Our scale contains 5 items (a VAS with both numbers and happy and sad faces and the endpoints 0=not having negative thoughts and 10=having negative thoughts). It asks children about both the presence and strength of five negative thoughts that are common in dentistry: uncontrollability, distrust of dentists, unpredictability, dangerousness, and pain related to dentistry [30,31]. We assessed the internal consistency of this scale by analyzing a sample from this study and a sample from another ongoing study; reliability is good (Cronbach alpha coefficient of .89, n=26).

Parental Self-Efficacy Questionnaire for Dental Anxiety (*P*-SEQ-DA), which is another measure constructed by our research team according to guidelines for constructing self-efficacy scales [32]. It comprises 12 items asking parents to evaluate their ability to support their children in dental situations (with the endpoints 0=no parental self-efficacy and 100=very high parental self-efficacy). A sample from this study showed the *P*-SEQ-DA to have good internal consistency (Cronbach alpha reliability coefficient of .89, n=26).

Injection phobia scale for children (IPSC), which measures changes in children’s degree of anxiety related to injection, consists of 18 items measuring the degree of fear associated with various situations associated to injection (scale 1-5, from no fear to high fear). The test has good reliability and validity [33,34].

In addition, participants and their parents responded to online questionnaires concerning the qualitative aspects of dental anxiety and ICBT with open answers, multiple choice questions, and VASs. We included questions evaluating the degree of satisfaction with the treatment. These questionnaires were delivered to the participants immediately after the treatment (week 13).

Procedure

The recruitment process began with an online screening comprising questionnaires about informed consent for primary caregivers and children and background information and inclusion and exclusion criteria for participation, as well as the CFSS-DS (child and parental), PG-BAT (child and parental), CNCD, and IPSC.

In the second step, parents received access to the Development and Well-Being Assessment (DAWBA) on the Internet and a personalized password. The DAWBA is a package of questionnaires and rating scales designed to generate the 10th revision of the International Statistical Classification of Diseases and Related Health Problems and DSM-IV psychiatric diagnoses for children in the age range of 5 to 17 years [35,36].

In the third step, parents were administered the K-SADS-PL over the phone to determine whether the child met the inclusion criteria for dental anxiety (including intraoral injection phobia) and other phobias (including general injection phobia, eg, vaccination). On the basis of the DAWBA and K-SADS-PL, a clinical psychologist determined if dental anxiety was the primary diagnosis and whether the participant met the inclusion criteria. If the child was over the age of 11 years, the psychologist conducted a short telephone interview with the child to ensure that the child had read the information about the project and was willing to participate. We conducted assessments before treatment (baseline), after treatment (12 weeks after the start of treatment), and at a 1-year follow-up. All assessments used online questionnaires except for the parent interview conducted with K-SADS-PL and questions about dental health staff and their acceptance of the treatment. These were administered by telephone interviews.

Intervention

The treatment manual comprised 12 modules (modules are distinct but interrelated units that ICBT is built upon) of guiding text (32,000 words), 18 worksheets, and 10 informational documents. Table 2 shows the content of modules. We wrote the manual in Swedish, developing it and adapting it to the Internet from a face-to-face CBT manual [14]. Four experienced clinical psychologists and one specialist in pediatric dentistry (coauthors) read the manual and gave continuous feedback during its development. Participants accessed treatment modules through a specially designed participant-secure platform for Internet-based psychological treatments.

Table 2. Contents of the modules.

Modules	Content
Intro (Modules 1-2)	
1-2	Coach psychoeducation; practical arrangements; home assignment; how to guide a child to elicit and reinforce behavioral change; rewarding strategies; and enhancing the child's self-efficacy
Exposure (Modules 3-11)	
3	Behavioral analyses; child psychoeducation and treatment rationale; goal setting
4	Constructing an exposure list and beginning exposure
5	Continued exposure (films and training package) and controlled breathing
6	Dentistry-related communication training; preparation for dental visit
7	Evaluation of dental visit; cognitive restructuring
8	Evaluation of ICBT[a] (so far); evaluation of exposure or treatment at a dental clinic; relaxation techniques
9	Pain and pain management education; fear, thoughts, and pain; focus shift and acceptance training
10	Problem solving and mindfulness training
11	Repetition; strategies for maintaining change and relapse prevention; letter to yourself
Diploma (Module 12)	
12	Relapse prevention plan; enhance your self-efficacy; diploma

[a] ICBT: Internet-based cognitive behavioral therapy.

Treatment Introduction

The first two modules targeted the coach, who was a parent or other person significant to the child and accepted to take primary responsibility for assisting the child during the treatment. From module 3 onward, the text addressed the children directly. Each module consisted of a number of tasks (answering questions and filling out worksheets) that the participant had to complete before moving on to the next module.

Figure 2 illustrates the components of the treatment. Early in treatment (module 2), we asked the coach to book a planning meeting (15 min) and a minimum of three dental appointments (30 min each) for exposure practice and dental care, to occur sometime between 6 weeks into the treatment and the end of treatment at 12 weeks. During treatment, the coach had access to an information sheet, available on the Internet platform that described the Internet treatment, basic aspects and rules of CBT, and the role of dental professionals during the Internet treatment. We asked coaches to send the information sheet to the dental clinics at which they had booked appointments. The dental care teams determined whether a dentist, dental hygienist, or dental assistant would take the main responsibility for exposure exercises at the dental clinic. The coach and children, together with their online psychologist, drafted suggestions concerning suitable exposure exercises that they then brought to the dental office for discussion.

Psychologist Contact

Each participant had a personal psychologist whom we introduced to the participant with a welcome letter at the beginning of treatment. Participants had continuous contact with their psychologist via a messaging system on the Internet platform. A psychologist guided the participants and their coaches during the 12 weeks of treatment. Two licensed psychologists provided this treatment. Both had a CBT qualification and experience in delivering CBT in pediatric dentistry. To increase treatment adherence and therapist competence, the psychologist with greater experience (8 years of experience as a therapist in pediatric dentistry) supervised the other psychologist on a weekly basis.

Homework

Each module ended with homework that contained both knowledge questions based on texts in the modules, as well as practical exercises such as exposures; registrations of, for example, negative thoughts related to dental care; relaxation exercises; and mindfulness. Exposure to dentistry-related video clips and audio files began from module 3. A practice package consisting of dental tools such as a dentist's mirror, dental probe, topical anesthetic, and cannula were sent home to the coach together with detailed instructions for use and safety (tool kit in Figure 2). Parents and children had to complete the assigned homework after each session before they were allowed to progress to the next module. The psychologist sent a message to the participant once a week. Messages consisted of feedback on homework assignment and answers to questions parents or children raised. If the assignments were not completed, reminders were sent to the participant. Exposures at the dental office started after module 6. The psychologists would provide feedback and support on homework assignments to participants within 36 hours on weekdays.

Figure 2. The components of the Internet-based treatment.

Relapse Prevention and Diploma

The treatment was ended by repetition, strategies for maintaining change, and relapse prevention. The participant has the possibility to print a diploma accessible in module 12.

The treatment manual for this study is available in Swedish and English and can be obtained by contacting the corresponding author.

Statistical Analyses

On the basis of an earlier study of CBT for pediatric dental anxiety [14], we expected a within-group Cohen *d* effect size of 0.80, which would require a sample size of 15 to achieve a power of 80% using a one-tailed test and alpha at .05.

The statistical analyses used Statistical package for the Social Sciences (SPSS) version 22 (IBM Corp). We analyzed data for the primary outcome measure (the PG-BAT, child and parental versions) and the other continuous measures (the CFSS-DS, child and parental versions; SEQ-SP; P-SEQ-DA; CNCS; and IPSC) using repeated measures analysis of variance. We also performed a separate sensitivity analysis for the primary outcome measure by putting missing data at the posttreatment and 1-year follow-up measurements (3 participants) with data from baseline assessments (assuming no change for these participants). Paired sample *t* tests compared continuous outcome measures from baseline to posttreatment, baseline to the 1-year follow-up, and posttreatment to the 1-year follow-up. Cochran Q test determined whether there were differences in the dichotomous dependent variables over time (from baseline to posttreatment to the 1-year follow-up). McNemar test explored if there were differences in the dichotomous dependent variables between two related groups (baseline to posttreatment, baseline to the 1-year follow-up, and posttreatment to the 1-year follow-up). Cohen *d*, based on pooled standard deviations, determined the effect sizes [37]. We evaluated clinically significant improvement on the basis of the primary outcome (level of avoidance behavior, as measured by the PG-BAT), setting a value of 12 or above as the cut-off value for clinically significant change. Values above 12 meant that the participant managed injection with local anesthesia, which is a necessary task in dentistry. Thus, we dichotomized the PG-BAT values based on a cut-off value of 12. This definition has been used in an earlier published study of face-to-face CBT for children and adolescents with dental phobia [14].

We also reported parents' responses to an online question, which asked if the participants had undergone and managed intraoral injection at a dental clinic after the ICBT.

Engagement, adherence, and compliance were used as a measure of feasibility: Engagement is the number of participants who began a module. Adherence is the number of participants who completed at least 75% of the modules (9/12 modules), and compliance is the number who completed the treatment program (all 12 modules). We considered a treatment module as complete when the participants had completed the home assignment and returned it to the psychologist on the Internet platform.

Results

Attrition and Dropout

Of 18 participants who began treatment, 2 discontinued participation in the study (Figure 1). In one case, the coach (mother) found the texts and home assignments too difficult to understand and manage. She declined participation after completing module 2. In the other case, the coach only performed the first assignment (module 1) and did not respond to reminders for continued participation. Another parent and her child dropped out after performing module 6. The parent stated that this was because of a crisis in their life situation and other priorities. However, the parent agreed to and partly completed the assessments at the posttreatment measurement, but they were not available for the 1-year follow-up.

Clinician Support and Adverse Effect

The average total clinician time per participant (including emails and telephone calls) was 5.4 hours (SD 2.3) which is approximately 30 min per week for each participant. We also asked participants and their parents about any negative or adverse effects of ICBT with a question in the online questionnaire they completed at the posttreatment assessment. None reported any negative effects.

Primary Outcome

Statistically significant changes in children's self-perceived ability to manage dental care procedures according to the PG-BAT, child version ($F_{2,28}$=14.1, P<.001) and in parental perceptions of the child's ability to manage dental care (not avoiding dental care), as measured by the PG-BAT, parental version ($F_{2,28}$=21.4, P<.001) occurred. There was a significant within-group improvement in both the child and the parental versions of the PG-BAT from baseline to posttreatment and from baseline to the 1-year follow-up (Table 3). No significant changes from posttreatment to the 1-year follow-up occurred. The sensitivity analyses showed no difference in significance levels from the original tests, which were missing data for three participants.

Secondary Outcome (Clinician Administered)

In the psychologist-administered telephone interviews with the parents, which used the K-SADS-PL (specific phobia section), there was a significant reduction over time in the proportion of participants with diagnosis dental anxiety (*P*=.001). We also found a significant reduction in the proportion of participants with a diagnosis of dental anxiety from baseline to posttreatment (*P*=.02) and from baseline to follow-up (*P*=.01). After treatment, 50% (7 /14) of participants no longer met the diagnostic criteria for dental anxiety and at the 1-year follow-up, 53% (8 /15). Corresponding reductions over time in injection phobia in the health care context (eg, vaccination) were also significant (*P*=.02). The proportion of participants with health care injection phobia reduced significantly from baseline to posttreatment (*P*=.008) but not from baseline to the 1-year-follow up (*P*=.06).

Table 3. Efficacy of cognitive behavioral therapy for children and adolescents with dental anxiety. Pre and post treatment and 1-year follow-up measures; *P* and *t* values; and effect sizes (according to Cohen *d*) and CIs.

Measures	Pre[a] (N=16) Mean (SD[b])	Post[c] (N=16) Mean (SD)	Follow-up[d] (N=15) Mean (SD)	Prepost			Prefollow-up		
				P value	*t* value	Effect sizes (95% CI)	*P* value	*t* value	Effect sizes (95% CI)
PG-BAT[e] child version	10.5 (1.0)	13.9 (2.9)	13.7 (2.8)	<.001	5.5	1.5 (0.7-2.3)	.001	4.1	1.4 (0.3-2.6)
PG-BAT parental version	9.3 (2.8)	13.7 (3)	13.7 (2.9)	<.001	4.9	1.5 (0.6-2.5)	<.001	4.8	1.6 (0.5-2.6)
CFSS-DS[f] child version	32.9 (10.2)	24.1 (6.8)	23.8 (6.4)	<.001	6.3	1.0 (0.5-1.6)	.006	3.2	1.1 (0.2-1.9)
CFSS-DS parental version	35.4 (10.1)	24.1 (6.3)	25.0 (4.8)	<.001	5.7	1.3 (0.7-2.0)	.001	4.5	1.3 (0.6-2.0)
SEQ-SP[g]	27.8 (8.3)	44.6 (6.9)	44.1 (11.1)	<.001	6.2	2.2 (0.8-3.6)	.001	4.3	1.66 (0.4-2.9)
P-SEQ-DA[h]	107.9 (13.3)	124.7 (7.2)	118.7 (11.6)	<.001	6.5	1.6 (0.9-2.2)	.02	2.6	0.9 (0.0-1.7)
CNCD[i]	24.0 (12.6)	8.0 (9.0)	11.4 (10.1)	<.001	5.4	1.7 (0.6-2.9)	.001	4.8	1.1 (0.5-1.8)
IPSC[j]	44.4 (14.3)	33.3 (10.9)	35.9 (12.1)	.001	4.05	0.9 (0.4-1.4)	.08	1.9	0.6 (−.05 to 1.3)

[a] Baseline measurement.

[b] SD: standard deviation.

[c] Posttreatment measurement (after 12 weeks of treatment).

[d] 1-year follow-up (1 year after posttreatment).

[e] PG-BAT: picture-guided behavior avoidance test.

[f] CFSS-DS: Children's Fear Survey Schedule-Dental Subscale.

[g] SEQ-SP: Self-Efficacy Questionnaire for Specific Phobias.

[h] P-SEQ-DA: Parental Self-Efficacy Questionnaire for Dental Anxiety.

[i] CNCD: children’s negative cognitions in dentistry.

[j] IPSC: injection phobia scale for children.

Other Outcomes (Self-Reported and Parent Reported)

There were significant changes in the participant’s self-perceived level of fear according to the CFSS-DS, child version ($F_{2,28}$=10.1, *P*<.001) and parental perceptions of the participant’s level of fear according to the CFSS-DS-P ($F_{2,28}$=18.6, *P*<.001). Similar results occurred in the child’s or adolescent’s level of self-efficacy, SEQ-SP ($F_{2,28}$=19.4, *P*<.001); parental self-efficacy, SEQ-DA-P ($F_{2,28}$=12.6, *P*<.001); the CNCD ($F_{2,28}$=19.4, *P*<.001); and for the general fear of injection ($F_{2,28}$=6.9, *P*=.01). There was a statistically significant within-group improvement from baseline to posttreatment and from baseline to the 1-year follow-up for all the continuous outcome measures except for the injection phobia scale, where the significant changes from baseline to posttreatment were not maintained at the 1-year follow-up (Table 3).

Effect Sizes

As Table 3 shows, large within-group effect sizes between baseline and posttreatment measurements and from baseline to the 1-year follow-up measurements occurred for both the primary measurements, level of avoidance (the PG-BAT), and other continuous measurements.

Clinically Significant Improvement

There was a significant difference in the proportion of participants who managed to pass the cut-off value over time (baseline to post to 1-year follow-up) for both child- and parent-reported PG-BATs (*P*=.001). At the posttreatment measurement, 56% (9/16) of parents and children reported that the participant could manage an intraoral injection. The corresponding proportion was 60% (9/15) at the 1-year follow-up. At baseline, all participants and their parents had reported that the participant could not manage local anesthesia. The improvement from baseline to posttreatment (*P*=.008 for children and *P*=.002 for parents) and from baseline to 1-year follow-up (*P*=.008 for both) was significant for both the child or adolescent and parental versions of the PG-BAT. Moreover, agreeing with the dichotomized results of the PG-BAT above, 60% (9/15) of parents reported at the 1-year follow-up that their child had managed intraoral injection at the dental clinic.

Feasibility and Acceptability

Measures of feasibility were engagement=18/18, adherence=14/18, and compliance=5/18. Children and their coaches completed, on average, 9.2 (SD 3.3) of the 12 treatment modules. Moreover, the completed measure (at posttreatment) rate was 90% (16/18). Measure of acceptance, that is, the average level of satisfaction with the treatment at posttreatment (from 1=quite dissatisfied to 4=very satisfied) for the participants was 3.3 (SD 0.6) and for the parents, 3.1 (SD 0.4). There were 15 children and 15 parents who were mostly satisfied or very satisfied (3 and 4 on the scale), whereas 1 child and 1 parent were indifferent (2 on the scale). Some parents and children, however, requested less extensive texts with fewer repetitions in the treatment manual. For all 15 children, the parents were able to book and attend at least three appointments at the dental clinic during the 12 weeks of ICBT. Participants visited a dental assistant, a hygienist, or a dentist during these visits.

Parents did not report (in the telephone interview at posttreatment) any major problems in terms of the dental personnel's acceptance of the Internet treatment. Parents reported that most of the dental professionals who met the child or adolescent during ICBT accepted the Internet treatment and showed interest in the treatment, meaning that they followed the information sheet that provided instructions for the staff. In two cases, however, coaches reported that the dental professionals were too cautious and avoided challenging exposures. According to these coaches, the dental professionals did not understand the importance of exposure and did not manage to help the children challenge their fear. Furthermore, two coaches felt uncomfortable with taking the time of the dental staff for training the child. We should also mention that, in three cases, the dental clinics meeting the participants initially asked for extra payment from our clinic for the extra costs of training the participants. However, when the clinic managers learned directly about the project from us, or via the parents, they dropped their requests for extra payment. We reminded these dental clinics that four visits, a total of 105 min, is not an unreasonable amount of time to train children who suffer from dental anxiety and that treatment of children with dental anxiety is a part of pediatric dentistry. In two cases, there was concern about getting too many participants from our project. We informed these clinics that no one clinic would receive too many such requests.

Discussion

Principal Findings

The aim of this study was to test the hypothesis that psychologist-guided ICBT improves school-aged children's and adolescents' ability to manage dental anxiety by (1) decreasing avoidance and affecting the phobia diagnosis and (2) decreasing the dental fear and increasing the target groups' self-efficacy. The study also aimed to examine the feasibility and acceptability of this novel treatment. The results show large within-group effect sizes (Cohen *d*) for the treatment, ranging from 0.9 to 2.2 for the outcome measures. Participants (8-15 years) improved their ability to cope with dental procedures, reduced their negative feelings and thoughts about dentistry, and increased their dentistry-related self-efficacy after the treatment. High levels of feasibility and acceptability (engagement, adherence, completed measures, and patient or parent satisfaction) were indicated by the results. Regarding compliance, the number of participants who completed all 12 modules was low, which indicates a need for reducing the number of modules.

Moreover, this study indicates that the treatment was acceptable from the perspective of dental health staff involved. Participating parents were able to book and attend at least three appointments at the dental clinic and meet a dental assistant, a hygienist, or a dentist. Parents overall reported high level of satisfaction concerning cooperation with the dental health staff that conducted the in vivo exposures.

Comparisons With Previous Work

Our results are in line with previous findings from a clinical trial testing a face-to-face CBT protocol for children and adolescents with dental anxiety [14]. The observed effect sizes of face-to-face CBT, however, were larger, ranging from 1.3 to 2.9 for the outcome measures. In the face-to-face CBT trial, participants managed more dental procedures in the BAT on average, and fewer participants met the diagnostic criteria for dentistry-related specific phobia compared with the participants in this ICBT study. This result might imply that ICBT is less effective than CBT, but these differences could also be because of differences in study design between the two studies, such as the differing age ranges for participants, different versions of the primary outcome measure, different level of severity of phobia for populations in these studies, and access to treatment at general or pediatric specialist dental clinics. ICBT combined with treatment at the specialist clinic could result in better treatment effect than ICBT combined with treatment in general dentistry. Thus, the results from these two studies are not entirely comparable. Due to small group sizes, we have not been able to conduct subanalyses that allow us to study the effect of age and symptom severity on the treatment efficacy. However, a meta-analyses conducted by another research team did not find age effects in results of exposure treatment (CBT) of child and adolescent anxiety [38].

Compared with self-help CBT for children with dental anxiety [21], our treatment is more resource-demanding (patient and parents are guided by a psychologist during 12 weeks). However, the self- help resource in dentistry we mentioned earlier and the material it consists of has not been evaluated in RCT studies. The effect of self-help CBT for patients with severe symptoms and the diagnosis of dentistry-related specific phobia is uncertain. There is some evidence in the literature that guided ICBT is more effective than unguided [39]. In this study, we wanted to build the new Internet-based treatment on the methods and material of the face-to-face psychologist-guided CBT program for pediatric dental phobia, which is evaluated in a RCT study [14] and use the benefits of guided ICBT.

One interesting finding of this trial was that the effect size for children's dentistry-related self-efficacy was as large as in the face-to-face study. This is a promising result as self-efficacy may be linked to clinically significant and sustainable behavioral change. In our earlier face-to-face CBT study, we hypothesized that differences in self-efficacy may explain why children in the treatment-as-usual group achieved, to a lesser extent, clinically significant behavior change [14]. Our results suggest that ICBT, like CBT, has the potential to affect self-efficacy. This is important as self-efficacy may be a mechanism of change in psychotherapy [40,41]. The self-efficacy effect could also be related to the film exposures included in both face-to-face therapy and ICBT. These film scenes were based on the principles of model learning and promoted processes that facilitate development of self-efficacy [40]. Studies have shown that vicarious exposure and film modeling can be as effective as or even more effective than live exposure when dealing with fear stimuli [42,43].

Another interesting finding was that the participants' fear of injection in general (in health care settings, eg, vaccination) decreased, as indicated by the clinical interview with parents and the injection phobia scale. The effect, however, was limited to changes from baseline to posttreatment measurements and was not maintained at the follow-up assessment.

In an earlier interview study with children and adolescent who had received CBT in dentistry, parents and children expressed uncertainty about the children's ability to manage injections outside of dental care (eg, vaccinations and blood sampling). Continued fear of injection in a health care setting could threaten improvement from CBT treatment in dentistry and enhance the risk of relapse to intraoral injection phobia [9]. Treatment module 12 contained recommendations for practices to help generalize the child's ability to receive injections in other health care settings. However, as mentioned in the results, only 5 of the participants completed and sent back the home assignments for module 12, which could mean that the remaining participants did not follow or use these recommendations. To address this, we should perhaps place the part of treatment dealing with general injection phobia in an earlier module and stress it more by incorporating planned exposures for injection training with health care staff.

The results of this study also agree with evaluations of evidence-based psychological treatments such as one-session CBT treatment for pediatric-specific phobias. In a study of one-session treatment, 49% of participants were free of a specific-phobia diagnosis at a 6-month follow-up [44] compared with 53% at the 1-year follow-up in our study. Others have reported similar results for ICBT in children with specific phobias, with approximately 50% of children at a 3-month follow-up free from a specific-phobia diagnosis [18,20].

Comparing one-session treatment and therapist-guided Internet treatment for spider phobia, both groups achieved clinically significant results according to the BAT. Approximately 70% achieved this effect at a 1-year follow-up [19]. Similar effects appear for ICBT in children with anxiety disorder, with 75% of participants reported as diagnosis-free at a 6-month follow-up [17]. The two booster sessions that participants were offered after treatment and before the follow-up may explain the higher percentage of diagnosis-free participants in the last study [20]. Implementing booster sessions in ICBT for pediatric dental anxiety may be a way to increase its treatment efficacy.

Limitations

This study is limited by its design and lack of a control group, which would make between-group comparisons possible. Participant improvements could hypothetically be a result of other factors than the treatment, such as time and maturation of the child; thus, the results should be interpreted with caution. However, CBT for specific phobias has shown good effects in numerous controlled studies [11,44]. On average, the participants in this study reported suffering from dental anxiety for 4 years before the study, which contradicts the supposition that this fear would disappear without treatment.

To take part in this study, a parent or coach needed to have the ability to manage the project information, apply for participation, and continuously support the participant during the project. Participants whose coach lacked these resources and abilities would have difficulty participating. As in several other ICBT studies, a high percentage of the coaches were well-educated, which may imply a limitation that parents with restricted resources and low education would struggle to access treatment.

Another limitation is that some of the measures used in this study were new and constructed in our research group. We had to construct these instruments because of a lack of suitable instruments for measuring behavioral and cognitive dimensions of pediatric dental phobia in the literature. These new measures need to be validated in larger studies.

Future Directions

Strengths of this study include a threefold measurement (child, coach, and psychologist assessments). Furthermore, the trial was conducted in a naturalistic dental care setting, which means that participants had the opportunity to experience exposure in vivo. Nevertheless, there is a need to modify and customize ICBT for participants and coaches who are less accustomed to reading texts, for instance by including short filmed lectures and fewer modules. Using ICBT to facilitate interaction between children or adolescents, coach or parent, and the medical or dental care staff is a novel approach implemented in this study. There is a potential to further develop this approach by increasing interaction with dental care staff. In this study, coaches brought short information sheets about ICBT to the staff they met. Providing access to brief ICBT courses for the dental care staff and supervision for them could increase treatment effect.

Another way to increase the dental health staff´s interaction could be to give them access to the Internet platform for the ICBT and allow the staff to take part; comment and give feedback to the participants during the whole ICBT. Although this may make the treatment more efficacious, there is a risk for making the treatment too demanding and time-consuming for the participating staff, which could decrease the staffs' acceptability. Studies that investigate the dental health staff acceptability for ICBT by qualitative interviews are important to conduct. Future studies should also test the ICBT for pediatric dental anxiety in RCTs.

Conclusions

If results of this open trial can be replicated in future controlled studies, there would be a potential for dental care systems to gain an effective evidence-based psychological treatment for children with dental anxiety—regardless of where they live or their access to specialist pediatric dental care —with relatively low personnel costs. While acknowledging the limitations of this study, we conclude that ICBT for dental anxiety in children and adolescents could be a feasible and efficacious treatment with the potential to increase accessibility to effective treatment.

Acknowledgments

The authors would like to thank the participating children or adolescents and their parents or coaches, dental assistant Katarina Edentorp Eriksson for help coordinating the project, psychologist Sara Ytterbrink for help constructing the Internet CBT platform and treating participants, the dental staff that met and trained the participants, the dental staff that acted in video clips included in the treatment program, Viggo and Lo Shahnavaz for enriching the treatment manual with a child's perspective through their feedback and comments and also for acting as participants in the video clips, professor Nils Lindefors for helping supervise the project, and the Internet-psychiatry unit at Stockholm County Council for giving participants access to the Internet treatment platform. They also thank the Mayflower Charity Foundation for Children, Karolinska Institutet, and the Stockholm County Council (SOF 16102563) for financial support of this project.

Authors' Contributions

SS, TH, EH, LR, and GD conceived and designed the experiments. SS and TH performed the experiments. SS, EH, VK, LR, and GD analyzed the data. SS, TH, EH, VK, LR, and GD contributed reagents or materials or analysis tools. SS wrote the paper. TH, EH, VK, LR, and GD reviewed and commented on the paper.

Conflicts of Interest

None declared.

References

1. Klingberg G, Broberg AG. Dental fear/anxiety and dental behaviour management problems in children and adolescents: a review of prevalence and concomitant psychological factors. Int J Paediatr Dent 2007 Nov;17(6):391-406. [doi: 10.1111/j.1365-263X.2007.00872.x] [Medline: 17935593]
2. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-IV-TR), 4th edition, text revision. Washington, DC: American Psychiatric Association; 2000:1-485.
3. Crego A, Carrillo-Díaz M, Armfield JM, Romero M. From public mental health to community oral health: the impact of dental anxiety and fear on dental status. Front Public Health 2014;2:16 [FREE Full text] [doi: 10.3389/fpubh.2014.00016] [Medline: 24616889]
4. Wide BU, Carlsson V, Westin M, Hakeberg M. Psychological treatment of dental anxiety among adults: a systematic review. Eur J Oral Sci 2013 Jun;121(3 Pt 2):225-234. [doi: 10.1111/eos.12032] [Medline: 23659254]
5. Roberts JF, Curzon ME, Koch G, Martens LC. Review: behaviour management techniques in paediatric dentistry. Eur Arch Paediatr Dent 2010 Aug;11(4):166-174. [Medline: 20840826]
6. Klingberg G, Andersson-Wenckert I, Grindefjord M, Lundin S, Ridell K, Tsilingaridis G, et al. Specialist paediatric dentistry in Sweden 2008 - a 25-year perspective. Int J Paediatr Dent 2010 Sep 01;20(5):313-321. [doi: 10.1111/j.1365-263X.2010.01057.x] [Medline: 20545788]
7. Mejàre IA, Klingberg G, Mowafi FK, Stecksén-Blicks C, Twetman SH, Tranæus SH. A systematic map of systematic reviews in pediatric dentistry--what do we really know? PLoS One 2015;10(2):e0117537 [FREE Full text] [doi: 10.1371/journal.pone.0117537] [Medline: 25706629]
8. Porritt J, Marshman Z, Rodd HD. Understanding children's dental anxiety and psychological approaches to its reduction. Int J Paediatr Dent 2012 Nov;22(6):397-405. [doi: 10.1111/j.1365-263X.2011.01208.x] [Medline: 22221141]
9. Shahnavaz S, Rutley S, Larsson K, Dahllöf G. Children and parents' experiences of cognitive behavioral therapy for dental anxiety--a qualitative study. Int J Paediatr Dent 2015 Sep;25(5):317-326. [doi: 10.1111/ipd.12181] [Medline: 26147012]
10. Ollendick TH, King NJ. Empirically supported treatments for children with phobic and anxiety disorders: current status. J Clin Child Psychol 1998 Jun;27(2):156-167. [doi: 10.1207/s15374424jccp2702_3] [Medline: 9648033]
11. Davis TE, May A, Whiting SE. Evidence-based treatment of anxiety and phobia in children and adolescents: current status and effects on the emotional response. Clin Psychol Rev 2011 Jun;31(4):592-602. [doi: 10.1016/j.cpr.2011.01.001] [Medline: 21482318]
12. Kvale G, Berggren U, Milgrom P. Dental fear in adults: a meta-analysis of behavioral interventions. Community Dent Oral Epidemiol 2004 Aug;32(4):250-264. [doi: 10.1111/j.1600-0528.2004.00146.x] [Medline: 15239776]
13. Haukebø K, Skaret E, Ost L, Raadal M, Berg E, Sundberg H, et al. One- vs. five-session treatment of dental phobia: a randomized controlled study. J Behav Ther Exp Psychiatry 2008 Sep;39(3):381-390. [doi: 10.1016/j.jbtep.2007.09.006] [Medline: 18005932]
14. Shahnavaz S, Hedman E, Grindefjord M, Reuterskiöld L, Dahllöf G. Cognitive behavioral therapy for children with dental anxiety. JDR Clin Trans Res 2016 Oct;1(3):234-243 [FREE Full text] [doi: 10.1177/2380084416661473]
15. Svensson L, Larsson Å, Öst L. How children experience brief-exposure treatment of specific phobias. J Clin Child Adolesc Psychol 2002 Mar;31(1):80-89. [doi: 10.1207/S15374424JCCP3101_10] [Medline: 11845654]
16. Hedman E, Ljótsson B, Lindefors N. Cognitive behavior therapy via the Internet: a systematic review of applications, clinical efficacy and cost-effectiveness. Expert Rev Pharmacoecon Outcomes Res 2012 Dec;12(6):745-764. [doi: 10.1586/erp.12.67] [Medline: 23252357]

17. Vigerland S, Lenhard F, Bonnert M, Lalouni M, Hedman E, Ahlen J, et al. Internet-delivered cognitive behavior therapy for children and adolescents: a systematic review and meta-analysis. Clin Psychol Rev 2016 Dec;50:1-10 [FREE Full text] [doi: 10.1016/j.cpr.2016.09.005] [Medline: 27668988]
18. Vigerland S, Thulin U, Ljótsson B, Svirsky L, Ost L, Lindefors N, et al. Internet-delivered CBT for children with specific phobia: a pilot study. Cogn Behav Ther 2013;42(4):303-314. [doi: 10.1080/16506073.2013.844201] [Medline: 24245708]
19. Andersson G, Waara J, Jonsson U, Malmaeus F, Carlbring P, Ost L. Internet-based self-help versus one-session exposure in the treatment of spider phobia: a randomized controlled trial. Cogn Behav Ther 2009;38(2):114-120. [doi: 10.1080/16506070902931326] [Medline: 20183690]
20. Vigerland S, Ljótsson B, Thulin U, Öst L, Andersson G, Serlachius E. Internet-delivered cognitive behavioural therapy for children with anxiety disorders: a randomised controlled trial. Behav Res Ther 2016 Jan;76:47-56 [FREE Full text] [doi: 10.1016/j.brat.2015.11.006] [Medline: 26649465]
21. Porritt J, Rodd H, Morgan A, Williams C, Gupta E, Kirby J, et al. Development and testing of a Cognitive Behavioural Therapy resource for children's dental anxiety. JDR Clin Trans Res 2017 Jan;2(1):23-37. [doi: 10.1177/2380084416673798]
22. Khanna MS, Kendall PC. Computer-assisted cognitive behavioral therapy for child anxiety: results of a randomized clinical trial. J Consult Clin Psychol 2010 Oct;78(5):737-745. [doi: 10.1037/a0019739] [Medline: 20873909]
23. Spence SH, Holmes JM, March S, Lipp OV. The feasibility and outcome of clinic plus internet delivery of cognitive-behavior therapy for childhood anxiety. J Consult Clin Psychol 2006 Jun;74(3):614-621. [doi: 10.1037/0022-006X.74.3.614] [Medline: 16822117]
24. Hedman E, Andersson G, Andersson E, Ljótsson B, Rück C, Asmundson GJ, et al. Internet-based cognitive-behavioural therapy for severe health anxiety: randomised controlled trial. Br J Psychiatry 2011 Mar;198(3):230-236 [FREE Full text] [doi: 10.1192/bjp.bp.110.086843] [Medline: 21357882]
25. Öst L, Skaret E, Hakeberg M, Raadal M, Haukebø K, Vika M. In: Öst L, Skaret E, editors. Cognitive Behaviour Therapy for Dental Phobia andAnxiety. Chichester, West Sussex, UK: John Wiley & Sons; 2013.
26. Davis T, Ollendick T, Öst L. In: Davis T, Ollendick T, Öst L, editors. Intensive One-Session Treatment of Specific Phobias. New York: Springer; 2012.
27. Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, et al. Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): initial reliability and validity data. J Am Acad Child Adolesc Psychiatry 1997 Jul;36(7):980-988. [doi: 10.1097/00004583-199707000-00021] [Medline: 9204677]
28. Klingberg G. Reliability and validity of the Swedish version of the Dental Subscale of the Children's Fear Survey Schedule, CFSS-DS. Acta Odontol Scand 1994 Aug;52(4):255-256. [Medline: 7985512]
29. Flatt N, King N. The Self-Efficacy Questionnaire for Phobic Situations (SEQ-SP): development and Psychometric Evaluation. Behav change 2012 Feb 22;26(02):141-152. [doi: 10.1375/bech.26.2.141]
30. Armfield JM, Slade GD, Spencer AJ. Cognitive vulnerability and dental fear. BMC Oral Health 2008 Jan 24;8:2 [FREE Full text] [doi: 10.1186/1472-6831-8-2] [Medline: 18218075]
31. Armfield JM. Towards a better understanding of dental anxiety and fear: cognitions vs. experiences. Eur J Oral Sci 2010 Jun;118(3):259-264. [doi: 10.1111/j.1600-0722.2010.00740.x] [Medline: 20572859]
32. Bandura A. Guide for Creating Self Efficacy Scale. In: Urdan TC, Pajares F, editors. SelfEfficacy Beliefs of Adolescents. Greenwich, CT: Information Age Pub., Inc; 2006:1-367.
33. Diva-portal. 2008. Injection phobi scale for children Internet URL: https://www.diva-portal.org/smash/get/diva2:199342/FULLTEXT01.pdf [accessed 2017-05-06] [WebCite Cache ID 6qFh4apcL]
34. Öst L, Hellström K, Kåver A. One versus five sessions of exposure in the treatment of injection phobia. Behav Ther 1992 Mar;23(2):263-281. [doi: 10.1016/S0005-7894(05)80385-5]
35. Goodman R, Ford T, Richards H, Gatward R, Meltzer H. The development and well-being assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. J Child Psychol Psychiatry 2000 Jul;41(5):645-655. [Medline: 10946756]
36. dawba. 2012. Development and Well-Being Assessment URL: http://www.dawba.com [accessed 2017-05-06] [WebCite Cache ID 6qFhcJo26]
37. Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [Medline: 19565683]
38. Bennett K, Manassis K, Walter SD, Cheung A, Wilansky-Traynor P, Diaz-Granados N, et al. Cognitive behavioral therapy age effects in child and adolescent anxiety: an individual patient data metaanalysis. Depress Anxiety 2013 Sep;30(9):829-841. [doi: 10.1002/da.22099] [Medline: 23658135]
39. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on Internet-based mental health interventions—A systematic review. Internet Interv 2014 Oct;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003]
40. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [Medline: 847061]
41. Gallagher MW, Payne LA, White KS, Shear KM, Woods SW, Gorman JM, et al. Mechanisms of change in cognitive behavioral therapy for panic disorder: the unique effects of self-efficacy and anxiety sensitivity. Behav Res Ther 2013 Nov;51(11):767-777 [FREE Full text] [doi: 10.1016/j.brat.2013.09.001] [Medline: 24095901]

42. Golkar A, Selbing I, Flygare O, Ohman A, Olsson A. Other people as means to a safe end: vicarious extinction blocks the return of learned fear. Psychol Sci 2013 Nov 01;24(11):2182-2190. [doi: 10.1177/0956797613489890] [Medline: 24022651]
43. Melamed BG, Weinstein D, Katin-Borland M, Hawes R. Reduction of fear-related dental management problems with use of filmed modeling. J Am Dent Assoc 1975 Apr;90(4):822-826. [Medline: 123932]
44. Ollendick TH, Ost L, Reuterskiöld L, Costa N, Cederlund R, Sirbu C, et al. One-session treatment of specific phobias in youth: a randomized clinical trial in the United States and Sweden. J Consult Clin Psychol 2009 Jun;77(3):504-516. [doi: 10.1037/a0015158] [Medline: 19485591]

Abbreviations

BAT: behavior avoidance test
CBT: cognitive behavioral therapy
CFSS-DS: Children's Fear Survey Schedule-Dental Subscale
CNCD: children’s negative cognitions in dentistry
DAWBA: Development and Well-Being Assessment
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders-4th edition
ICBT: Internet-based cognitive behavioral therapy
IPSC: injection phobia scale for children
K-SADS-PL: Kiddie-Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime version
PG-BAT: picture-guided behavior avoidance test
P-SEQ-DA: Parental Self-Efficacy Questionnaire for Dental Anxiety
RCT: randomized controlled trial
SEQ-SP: Self-Efficacy Questionnaire for Specific Phobias
SD: standard deviation
VAS: visual analog scale

Edited by G Eysenbach; submitted 08.05.17; peer-reviewed by N Alberts, J Dudeney, J Apolinário-Hagen; comments to author 15.06.17; revised version received 30.09.17; accepted 02.11.17; published 22.01.18.

Please cite as:
Shahnavaz S, Hedman-Lagerlöf E, Hasselblad T, Reuterskiöld L, Kaldo V, Dahllöf G
Internet-Based Cognitive Behavioral Therapy for Children and Adolescents With Dental Anxiety: Open Trial
J Med Internet Res 2018;20(1):e12
URL: http://www.jmir.org/2018/1/e12/
doi:10.2196/jmir.7803
PMID:29358158

Original Paper

Implementing a Mobile Health System to Integrate the Treatment of Addiction Into Primary Care: A Hybrid Implementation-Effectiveness Study

Andrew Quanbeck[1,2], PhD; David H Gustafson[2], PhD; Lisa A Marsch[3], PhD; Ming-Yuan Chih[4], PhD; Rachel Kornfield[5], MA; Fiona McTavish[2], MSc; Roberta Johnson[1], MA, MEd; Randall T Brown[1], MD, PhD; Marie-Louise Mares[6], PhD; Dhavan V Shah[5], PhD

[1]Department of Family Medicine and Community Health, University of Wisconsin - Madison, Madison, WI, United States
[2]Center for Health Enhancement Systems Studies, University of Wisconsin - Madison, Madison, WI, United States
[3]Center for Technology and Behavioral Health, Dartmouth College, Lebanon, NH, United States
[4]College of Health Sciences, University of Kentucky, Lexington, KY, United States
[5]School of Journalism and Mass Communications, University of Wisconsin - Madison, Madison, WI, United States
[6]College of Letters and Science, University of Wisconsin - Madison, Madison, WI, United States

Corresponding Author:
Andrew Quanbeck, PhD
Department of Family Medicine and Community Health
University of Wisconsin - Madison
1100 Delaplaine Ct
Madison, WI, 53715
United States
Phone: 1 608 263 4550
Fax: 1 608 263 5813
Email: arquanbe@wisc.edu

Abstract

Background: Despite the near ubiquity of mobile phones, little research has been conducted on the implementation of mobile health (mHealth) apps to treat patients in primary care. Although primary care clinicians routinely treat chronic conditions such as asthma and diabetes, they rarely treat addiction, a common chronic condition. Instead, addiction is most often treated in the US health care system, if it is treated at all, in a separate behavioral health system. mHealth could help integrate addiction treatment in primary care.

Objective: The objective of this paper was to report the effects of implementing an mHealth system for addiction in primary care on both patients and clinicians.

Methods: In this implementation research trial, an evidence-based mHealth system named Seva was introduced sequentially over 36 months to a maximum of 100 patients with substance use disorders (SUDs) in each of three federally qualified health centers (FQHCs; primary care clinics that serve patients regardless of their ability to pay). This paper reports on patient and clinician outcomes organized according to the Reach, Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM) framework.

Results: The outcomes according to the RE-AIM framework are as follows: Reach—Seva reached 8.31% (268/3226) of appropriate patients. Reach was limited by our ability to pay for phones and data plans for a maximum of 100 patients per clinic. Effectiveness—Patients who were given Seva had significant improvements in their risky drinking days (44% reduction, (0.7-1.25)/1.25, *P*=.04), illicit drug-use days (34% reduction, (2.14-3.22)/3.22, *P*=.01), quality of life, human immunodeficiency virus screening rates, and number of hospitalizations. Through Seva, patients also provided peer support to one another in ways that are novel in primary care settings. Adoption—Patients sustained high levels of Seva use—between 53% and 60% of the patients at the 3 sites accessed Seva during the last week of the 12-month implementation period. Among clinicians, use of the technology was less robust than use by patients, with only a handful of clinicians using Seva in each clinic and behavioral health providers making most referrals to Seva in 2 of the 3 clinics. Implementation—At 2 sites, implementation plans were realized successfully; they were delayed in the third. Maintenance—Use of Seva dropped when grant funding stopped paying for the mobile phones and data plans. Two of the 3 clinics wanted to maintain the use of Seva, but they struggled to find funding to support this.

Conclusions: Implementing an mHealth system can improve care among primary care patients with SUDs, and patients using the system can support one another in their recovery. Among clinicians, however, implementation requires figuring out how information from the mHealth system will be used and making mHealth data available in the electronic health (eHealth) record. In addition, paying for an mHealth system remains a challenge.

(J Med Internet Res 2018;20(1):e37) doi:10.2196/jmir.8928

KEYWORDS

mobile health; mHealth; evidence-based practice; behavioral medicine

Introduction

The Use of mHealth in Primary Care

Despite the near ubiquity of mobile phones, little systematic research has been conducted on the use or implementation of mobile health (mHealth) technology in managing chronic health conditions in primary care. Although primary care patients use mHealth apps, their use is generally haphazard and self-selected [1,2]. The mHealth apps available to the public vary greatly in quality, and problems such as software bugs, poor design, and limited technical support are common [3]. Perhaps not surprisingly, given these problems, the majority of health-related apps are used only once [4]. Primary care clinicians know very little about the mHealth systems their patients are using, and most clinicians receive no health-related information from mHealth systems [1,5]. The literature contains numerous pilot studies and descriptions of mHealth systems [6] but only a few rigorous studies about the use of mHealth in primary care [5,7]. Implementation research studies in which primary care clinics enroll cohorts of patients using the same mHealth system in an orchestrated fashion are particularly lacking. Thus, the role mHealth can play in the US primary care system remains largely unknown [8].

mHealth Interventions for Addiction Treatment

Although hundreds of mobile phone apps for addiction treatment are available commercially, most of these apps have not been evaluated in the peer-reviewed literature [9]. The small number of apps that have been evaluated constitute a growing body of evidence supporting the effectiveness of mHealth in treating addiction [10-13]. Most of this evidence relates to self-help interventions [10,11,13] and to texting-based monitoring and reminder systems [9,12]. Little evidence relates to comprehensive mHealth systems for addiction, which have the strongest theoretical base and the most long-lasting effects, or to mHealth interventions for addiction integrated into patients' recovery and health plans [9].

Barriers to the Integration of Behavioral Health in Primary Care

In this paper, we report the results of an implementation research trial funded by the National Institutes of Health–National Institute on Drug Abuse. The trial aimed to integrate behavioral health treatment into primary care. We focused on one aspect of behavioral health—addiction—that presents considerable barriers to integration, such as the inability to bill for services, mental health stigma, and primary care physicians being ill-prepared to treat behavioral health problems [14,15]. Primary care operates under productivity guidelines that limit the time clinicians can spend with patients, whereas addiction treatment typically involves frequent counseling sessions. Medication is crucial in primary care, but it has a comparatively short history in addiction treatment [16]. Primary care focuses on chronic conditions, such as diabetes and hypertension, and practitioners expect that patients' adherence to treatment will vary over time. Behaviorial health has only recently begun to view addiction as a chronic condition [16-18]. Lapses have often resulted in discharge from treatment. Primary care treats patients one-on-one, whereas behavioral health often organizes patients into groups for treatment [19]. Financing models and information technology (IT) infrastructure also differ greatly between the two systems of care [17]. Finally, patients with addictions often have elevated anxiety and present frustrating behaviors to providers, such as frequently missing appointments [20].

In this research, our premise was that mHealth could ease the integration of addiction treatment into primary care. We proposed to examine how implementing an evidence-based mHealth system for addiction could be useful to both patients and clinical staff in real-world primary care settings. The mHealth system used in the study is named *Seva*, a Sanskrit word meaning "selfless caring." Its key components were previously proven effective in carefully controlled patient-level randomized clinical trials [21,22]. In a randomized clinical trial conducted in patients with alcohol use disorder leaving 90-day residential care, the intervention comprising the backbone of Seva reduced risky drinking days by 57% [21] and increased retention in treatment by 77% [23] compared with patients in the control group.

Purpose of This Study

The study reported here sought answers to 3 broad research questions:

1. How can Seva be implemented in primary care settings efficiently and effectively?
2. To what extent do patients and staff accept and use Seva?
3. How does Seva affect clinical care for patients and staff?

Methods

Study Design

Because the study tested both clinical and implementation interventions, it is considered to be a hybrid type 2 effectiveness-implementation study [24]. Details of the study protocol, the theoretical foundations of the implementation strategy, and a description of Seva were published previously [25]. This paper reports quantitative and observational results. Selected qualitative results were reported separately [26].

We made Seva available to up to 100 patients at each of 3 federally qualified health centers (FQHCs) across the United States of America. FQHCs are federally funded primary care clinics that serve mainly low-income patients. As a condition of funding, FQHCs must provide access to behavioral health services. Thus, FQHCs are in the vanguard of clinics in the United States trying to integrate behavioral health into primary care. FQHCs also serve very vulnerable patients—many in poverty and suffering from addiction. At each site, patients were enrolled over a 12-month period. After enrollment, patients had access to Seva for 12 months.

Because the focus of the study was implementation and not patient outcomes *per se*, the study did not randomize patients. Instead, clinicians were free to enroll any patients from their substance-using populations whom they thought might benefit from Seva based on their clinical judgment. We used the Reach, Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM) framework to organize the evaluation [27]. The RE-AIM framework is a predominant evaluation framework in implementation research studies.

Ethics

The study protocol was approved by the Medical Sciences Institutional Review Board at the University of Wisconsin–Madison (2012-0937-CP020) and is registered at ClinicalTrials.gov (NCT01963234).

Clinic Recruitment

Clinics were recruited in partnership with the National Association of Community Health Centers. We recruited FQHCs with established electronic health (eHealth) records to understand how Seva relates to existing clinic technology. In selecting sites, we aimed to achieve geographic reach, diversity in patient populations, and differences in organizational structures to better understand how environmental and structural factors might affect implementation. From a pool of approximately 1100 FQHCs nationally, we selected an FQHC affiliated with the University of Wisconsin as a pilot site; a relatively small, rural, freestanding FQHC with integrated behavioral health services (including addiction treatment) as a second site; and an urban FQHC that largely serves a minority population as our third site.

Clinician and Patient Recruitment

A limited number of clinician subjects consented at each site. These staff subjects worked with the research team to integrate Seva into clinical workflows during the preimplementation phase and subsequently identified, recruited, enrolled, and trained patients to use Seva and monitored patients’ use of the system.

Authorized clinicians identified potential subjects through the electronic health record (EHR). On the basis of their clinical judgment, clinicians were free to enroll patients from their substance-using populations whom they thought might benefit from Seva. The patients had to meet the following inclusion criteria: (1) aged 18 years or older, (2) meet the criteria for substance use disorder (SUD) as per the Severity of Dependence Scale, (3) have no current psychotic disorder severe enough to prevent participation, (4) have no acute medical problem requiring immediate inpatient treatment, (5) are willing to use Seva, and (6) could understand and sign a consent form in English. If a patient was incarcerated during the study, his or her participation was stopped. If he or she was still interested in participating in the study when released from jail, the person was able to rejoin the study. Patient participation was voluntary. Patient subjects were excluded if (1) their condition warranted inpatient detoxification until they were well enough to participate and (2) they were unable to understand and complete the informed consent. Couples were not recruited to the study at the same time to avoid dyadic conflict.

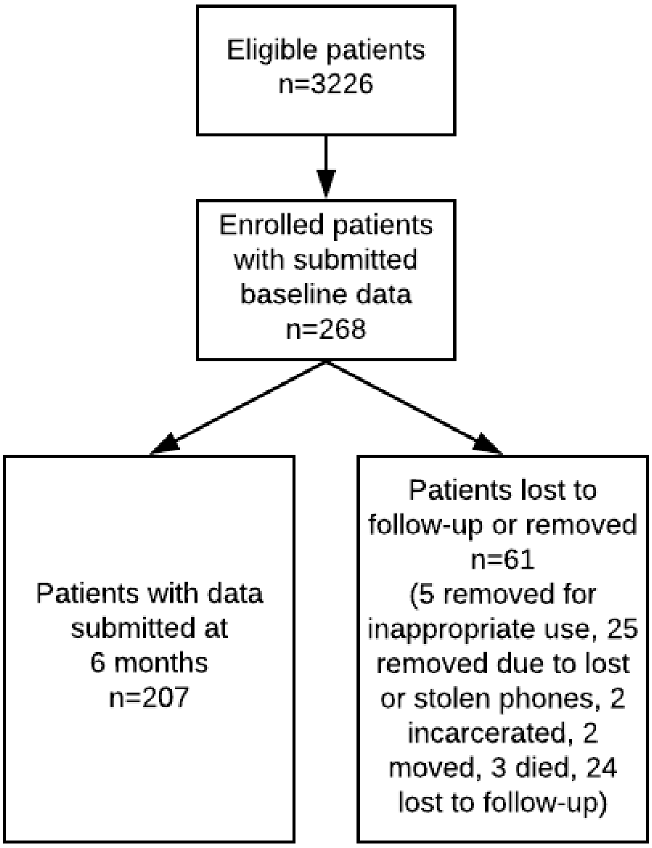
During a clinic appointment, the clinician asked an eligible patient if he or she was interested in learning more about the study. If the patient was interested and gave permission, the clinician notified the site coordinator that a patient was interested in hearing about the project. In the privacy of a clinic exam room or office, the site coordinator explained the study, its benefits, and potential risks of participation. The site coordinator also answered any questions the patient (subject) had. If the patient was interested in participating, he or she was asked to complete the informed consent.

The flow of patients through the study is shown in Figure 1.

Training Patients on Seva

The site coordinator trained participants in person. Participants could download Seva if they had their own mobile phone. If they did not, the site coordinator gave participants a phone and basic instructions on how to use it. Then he or she explained and showed participants how to use the different services of Seva. Participants were encouraged to ask questions and were given a toll-free number to call the Seva technical support line if they had any questions. Each participant also received a user guide that showed all Seva services and explained how to use them with easy text and graphic instructions.

Figure 1. Participant flowchart.



Clinic Rollout Process

Each site designated a change leader, a clinical leader who was the point of contact with the implementation coach and coordinated implementation activities; the site coordinator; and a change team of 4 to 8 clinical and/or administrative staff members who helped make the organizational changes that were necessary to implement Seva. The coach made an initial site visit during the 4-month period of implementation preparation to create a welcoming environment for Seva. During this visit, the coach conducted a walk-through exercise with the change team members (an exercise in which employees experience clinic processes as patients do); a workflow assessment using flowcharting; and a technical assessment of data to be gathered and procedures needed to conduct the study. Through this preparation period, the coach worked with the change leader to ensure that pretest data were collected, Seva was demonstrated to clinical teams who might serve as referral sources, and barriers to implementation were identified and rectified.

mHealth Intervention

Seva offered patients a discussion board used by the patients in the study; interactive modules to teach problem-solving, self-regulation, and other skills; tools for coping with cravings and high-risk situations (eg, relaxation exercises, strategies from cognitive behavioral therapy, links to local 12-step meetings); and health tracking. For clinicians, Seva provided a Web portal with a Clinician Report containing longitudinal information generated by patients' self-reported data about their substance use and well-being (eg, sleep, depression).

Implementation Strategy

We created a detailed implementation plan tailored to each clinic. The implementation strategy involved 4 phases at each site, which were (1) initiate (bring key clinical staff together for training), (2) prepare (assure a welcoming environment for Seva), (3) improve (conduct rapid-cycle tests of ideas from the previous stages), and (4) implement (use, monitor, and sustain the technology). An organizational coach (a member of the research team experienced in coaching) helped clinics implement Seva, starting with an in-person planning and kick-off meeting. This meeting was to be followed by monthly phone calls. Implementation plans consisted of a list of scheduled activities, such as when recruitment would start and end, when patients would be trained to use Seva, and so on. The implementation plans were informed by baseline assessments of readiness for implementation, sustainability potential, and technology acceptance from the perspective of clinic staff members who worked on the implementation plan [28-30].

We expected that Seva would enable primary care physicians to better manage addiction in their patients, much as they manage chronic conditions such as diabetes and asthma, by making information about a patient's recovery readily available. Physicians and other clinicians who referred a patient to behavioral health could see through Seva how these referrals worked out. Seva would also improve outcomes by making support, information, and skills training available to patients almost anywhere and anytime.

Outcomes

The measures used in this study are summarized in Table 1.

Table 1. List of measures.

Domain and measure	Data source(s)
Reach	
Number of Seva patients (eligible, enrolled)	EHR[a], patient surveys
Characteristics of participating patients	Patient surveys
Effectiveness	
Substance use	Patient surveys
QoL[b]	Patient surveys
Health care utilization (hospitalizations, ER[c] visits, specialty addiction treatment)	Patient surveys
HIV testing rates	Patient surveys
HIV risk behaviors	Patient surveys
Adoption	
Characteristics of participating clinics	Clinic administrative data
Use of Seva by staff (including referrals)	Seva server files, referral tracking logs maintained by clinic staff
Use of Seva by patients	Seva server files
Implementation	
Stages of Implementation Completion	Project tracking logs maintained by research team
Implementation and operating costs	Observation and interviews of clinic staff; project administrative data
Maintenance	
6-month follow-up on effectiveness measures	Patient surveys
Patient use of Seva at 12 months	Seva server files

[a] EHR: electronic health record.

[b] QoL: quality of life.

[c] ER: emergency room.

To assess *Reach*, we examined the characteristics of the participating sites and the patients they serve. We examined how many patients were eligible to use Seva in each clinic versus how many enrolled in the study over the enrollment period.

To assess *Effectiveness*, we considered patients' substance use, quality of life (QoL), health care utilization, human immunodeficiency virus (HIV) risk behaviors, and HIV testing rates through self-reported data collected in surveys administered in person at baseline and by phone at 6 months. Risky drinking days were the number of days in the past 30 days on which, within a period of 2 hours, women consumed more than 3 standard drinks and men more than 4 standard drinks, corresponding to the National Institute on Alcohol Abuse and Alcoholism's definition of binge drinking [31]. Illicit drug use days were defined as the number of days in the past 30 days on which participants used any illicit drug. QoL was measured using the PROMIS Global Health Scale [32].

Effectiveness outcomes are reported for differences between baseline and 6 months. To handle skew distributions, nonparametric related sample tests (Wilcoxon signed rank tests for continuous and ordinal variables and McNemar tests for binary outcomes) were conducted. *P* values (2-tailed) and effect sizes were reported.

We assessed *Adoption* at both the patient and clinic level. For patients, we assessed weekly rates of Seva use, with *use* defined as a patient accessing any part of Seva beyond the home page during a given week in the 12-month period when patients' phones and service plans were paid for. We defined *clinician use* in two ways—by the number of log-ins by clinicians per week to the website where Seva data were available and by the number of referrals to Seva. We also tracked the total number of Seva patients enrolled and whether medical providers (doctors of medicine [MDs], residents, physician assistants [PAs], nurse practitioners [NPs], and nurses or behavioral health providers) referred patients to Seva. To assess *Implementation*, we used the Stages of Implementation Completion model [33]. The stages of implementation completion are (1) engagement; (2) consideration of feasibility; (3) readiness planning; (4) staff hired and trained; (5) fidelity monitoring processes in place; (6) services and consultation begin; (7) ongoing services, consultation, fidelity monitoring, and feedback; and (8) competency. We planned to assess *Maintenance* in a follow-up phone survey 6 months after each patient's 12-month intervention period ended, but this plan was abandoned, as discussed below.

Results

Reach

We selected an FQHC in Madison, WI, affiliated with the University of Wisconsin-Madison as our pilot implementation site; a relatively small, rural, freestanding FQHC with integrated behavioral health services, including addiction treatment, in Missoula, MT, was our second site; and an urban FQHC in the Bronx, NY, that serves a largely minority population was our third site (Table 2). In all 3 clinics, a total of 3226 patients were deemed clinically appropriate to use Seva in the 12 months corresponding to the implementation period. This number of appropriate patients represents all patients with a SUD diagnostic code in the EHR. With 268 patients enrolled, the intervention reached approximately 8.31% (268/3226) of patients with substance use issues at these 3 clinics. The reach of Seva to the target population was limited by our ability to pay for phones and data plans for a maximum of 100 patients per clinic. The racial and ethnic composition of patients in the study departs somewhat from those of US adults 18 or older estimated by the Substance Abuse and Mental Health Services Administration to meet the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) criteria for alcohol dependence and abuse, with study participants being 67.9% (182/268) white versus 55% nationally; 25.0% (67/268) African American/Black versus 16% nationally; 11.6% (31/268) Other versus 10% nationally; and 14.2% (38/268) Hispanic/Latino versus 20% nationally [34]. At least one clinic indicated more patients could have been enrolled.

Effectiveness

Tables 3 and 4 show changes in patient outcomes reported from baseline to 6 months. Of 268 enrolled patients, 207 (77.2%) were included in this analysis. Effect sizes were calculated per Cohen [35].

In the substance use domain, significant reductions were observed for the number of risky drinking days (the primary outcome in the clinical trial preceding this implementation study), which declined by 44% [(0.7-1.25)/1.25] from baseline to 6 months and illicit drug-use days, which declined by 34% [(2.14-3.22)/3.22]. Two of the three abstinence outcomes also showed significant improvements (any illicit drug use and/or any drink or drug). Significant effects were found for two of the three QoL scores (overall QoL and mental health). Table 3 also shows a significant reduction in hospitalizations and a trend toward fewer emergency room (ER) visits. Table 4 shows an increase in HIV screening rates. Change in the rates of HIV risk behaviors (eg, condom use) and receiving other addiction treatments appeared to be nonsignificant.

Post hoc analyses assessed the relationship between the extent of Seva use and our study outcomes at 6 months. For each outcome, we used linear or logistic regression, controlling for the value of the outcome at baseline. System use was operationalized as (1) the total number of calendar days during the first 6 months on which individuals used Seva (going beyond the main menu), and (2) the total number of Seva pages viewed in the first 6 months (excluding the main menu). These system use measures were natural log transformed to reduce skewness. We found that participants who used Seva on more days in the first 6 months showed a significant increase in alcohol abstinence (*P*=.02), and participants who loaded more Seva pages showed a significant increase in overall abstinence from both alcohol and drugs (*P*=.01), as well as reduced HIV risk behaviors (*P*=.02). We found no significant associations between system use and risky drinking days, illicit drug use days, health care utilization, or QoL.

Whereas Seva's quantitative results show promise in helping primary care patients remain abstinent, other important effects can be best appreciated by directly observing how mHealth affected patient care. An exchange on the Seva discussion board in the Bronx (Site 3) illustrates how the system helped patients support one another. Figure 2 shows an unedited excerpt of this exchange—only names have been changed. The exchange—which occurred over a 2-hour period beginning at 4:41 AM, well before the clinic was open—illustrates how patients struggling with addiction can support one another, in real time, outside the clinic, using mHealth. As the exchange highlights, the network of patients in the Bronx even took to referring to themselves as the "Seva family."

Adoption

Mobile health apps generally have low levels of continued use—approximately 80% are abandoned after only 2 weeks [36]. In this context, all 3 sites showed high levels of sustained patient use (Figure 3). At the start of the study, rates of patient use—defined as accessing any part of Seva beyond the home page during each week of the 12-month implementation period—at the 3 sites ranged from 94% (90/96) to 99% (69/70). Rates of use declined, but slowly, with rates at 12 months ranging from 53% (41/78) to 60% (39/65) across the 3 sites, which mirrors the patient retention rate of 57.6% at 8 months in the randomized trial of an earlier iteration of Seva [21].

Clinician adoption of Seva was less robust than patient adoption for two main reasons. At each clinic, clinicians worried about being responsible for data available from Seva. For example, clinicians were concerned that a patient might express suicidal thoughts in a discussion post and they would miss it. To address this concern, the clinical staff at each clinic wanted one staff member to lead the implementation and operation of Seva for the clinic. This job included monitoring patient Seva activity for the clinic and alerting clinicians of significant changes in patient status. At 2 of the 3 sites, members of the behavioral health department led the implementation and operation of Seva. At the other site, the director of innovation led implementation. Not all FQHCs have such a position. At one of the sites, the behavioral health provider who led the implementation of Seva left the clinic near the end of the patient enrollment period. Members of the research team increased their involvement at the site to pick up the slack, but staff turnover remained a vexing issue.

Table 2. Baseline characteristics of participating clinics and patients.

Characteristics	Site 1 (Madison, WI)	Site 2 (Missoula, MT)	Site 3 (Bronx, NY)
Clinic characteristics			
Insurance of patients (%)			
Insured			
Medicare	8.7	11.6	11.1
Medicaid	56.3	23.1	46.5
Private/other	18.5	19.4	28.9
Not insured	16.5	45.9	13.5
eHealth[a] records	Epic	eClinicalworks	Epic
Services offered	Primary care and mental health	Primary care, mental health, and addiction	Primary care and mental health
PCMH[b] designation	Level 3 (2011)	Level 3 (2014)	Level 3 (2014)
Patient characteristics			
Number of eligible SUD[c] patients	1189	961	1076
Patients enrolled in Seva	97	100	71
Age (years)			
Range	21-64	21-66	22-64
Mean (standard deviation)	41.61 (10.95)	42.53 (10.24)	42.66 (11.78)
Gender n, (%)			
Female	52 (54)	40 (40)	36 (51)
Drug of choice, n (%)			
Alcohol	34 (35)	44 (44)	27 (38)
Opiates	31 (32)	14 (14)	8 (11)
Crack cocaine	9 (9)	3 (3)	11 (16)
Marijuana	1 (1)	4 (4)	16 (23)
Methamphetamine	0 (0)	15 (15)	1 (1)
Multiple drugs	22 (23)	20 (20)	8 (11)
Ethnicity n (%)			
Hispanic/Latino	1 (1)	2 (2)	35 (49)
Race, n (%) [d]			
White	68 (70)	90 (90)	24 (33)
African American/Black	30 (31)	2 (2)	35 (50)
American Indian or Alaskan Native	4 (4)	8 (8)	1 (2)
Asian or Pacific Islander	0 (0)	1 (1)	0 (0)
Other	0 (0)	2 (2)	15 (21)

[a] eHealth: electronic health.

[b] PCMH: patient-centered medical home. Three levels of recognition exist, based on practice sites meeting six standards. Level 3 clinics have the best adherence to the standards.

[c] SUD: substance use disorder.

[d] Percentages do not add to 100 because patients could select more than one race.

Table 3. Effectiveness results, continuous patient outcomes

Measures	Sample size N	Baseline Mean (SD)	6 months Mean (SD)	Z[a,b] (*P* value)	Effect size *d*[b,c]
Substance use in last 30 days					
Any drinking days	207	2.53 (6.01)	1.67 (4.69)	-2.304 (.02)	-0.228
Risky drinking days[d]	207	1.25 (3.78)	0.70 (2.58)	-2.008 (.4)	-0.199
Illicit drug-use days	206	3.22 (7.57)	2.14 (6.55)	-2.499 (.01)	-0.248
QoL [e]					
Overall QoL	202	28.47 (6.46)	30.03 (7.11)	3.653 (<.001)	0.370
Physical subscale QoL	206	13.20 (3.01)	13.48 (3.11)	1.682 (.09)	0.167
Mental subscale QoL	204	9.75 (2.99)	10.77 (3.50)	3.892 (<.001)	0.393
Health care utilization in last 6 months					
No. of hospitalizations[d]	207	0.43 (1.03)	0.22 (0.65)	-3.357 (.001)	-0.335
No. of ER[f] visits[d]	207	1.10 (2.79)	0.75 (1.31)	-1.911 (.06)	-0.189

[a] Z, provided in the Wilcoxon sign test, is the standard normal distributed *Z*-value used to test the significance between outcomes reported at two time points (eg, pretest vs 6 months).

[b] For *Z* and *d* values, negatives mean decreases and positives mean increases in values from baseline to 6 months.

[c] Calculated from effect size *d*. On the basis of Cohen (1988) effect size, small: *d*=0.2, medium: *d*=0.5, large: *d*=0.8.

[d] Risky drinking days, hospitalizations, and ER visits: Those who reported no such events were coded with zero in the number of days of these events.

[e] QoL: quality of life.

[f] ER: emergency room.

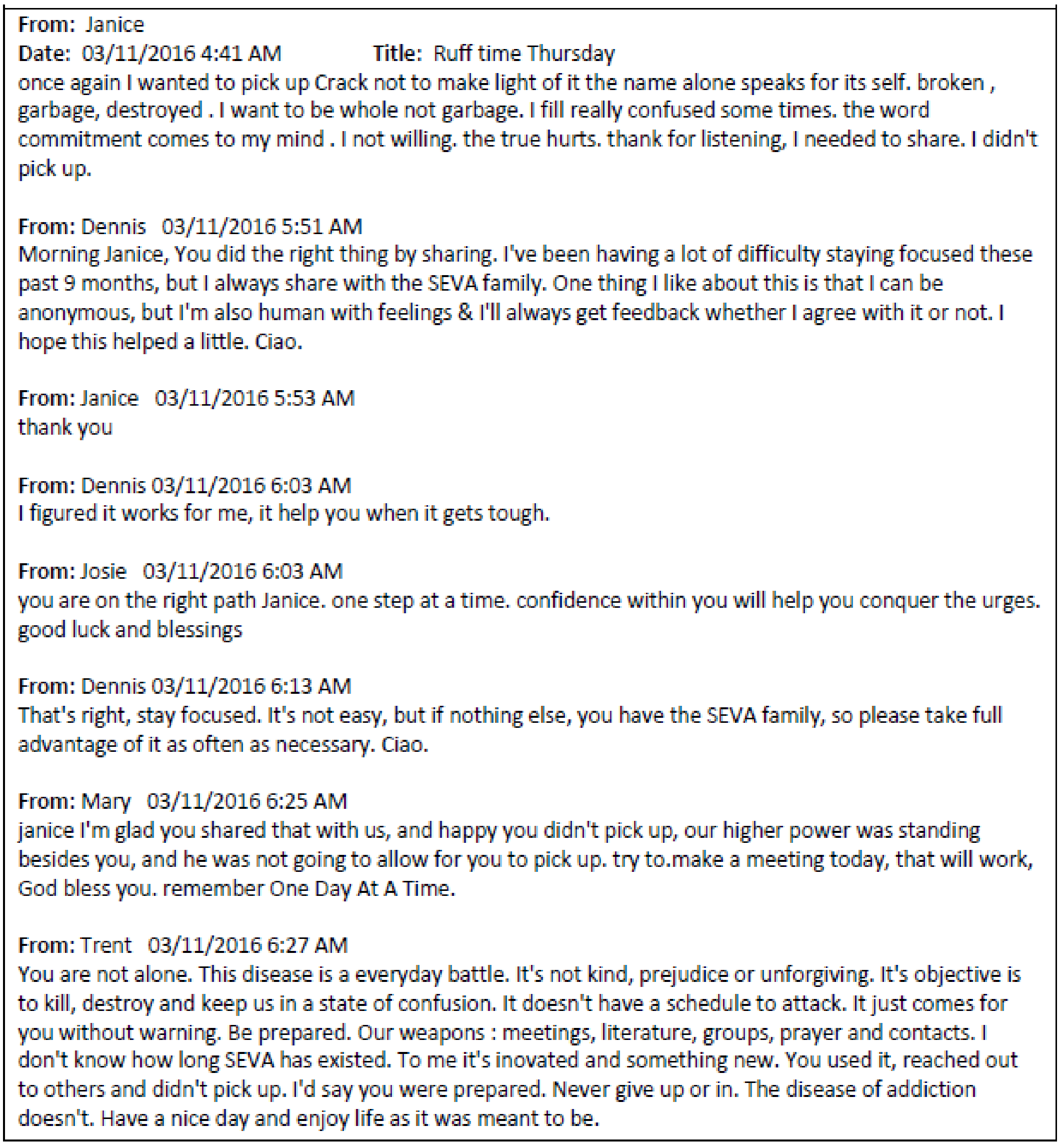
Table 4. Effectiveness results, dichotomized patient outcomes

Measures	Sample size N	Baseline n (%)	6 months n (%)	Chi-square (*P* value)	Odds ratio[a]
Substance use in last 30 days					
Any drink (Yes)	207	64 (30.9)	51 (24.6)	3.2 (.07)	0.552
Illicit drug use (Yes)	206	63 (30.6)	36 (17.5)	14.38 (<.001)	0.270
Any drink or drug (Yes)	206	97 (47.1)	69 (33.5)	12.57 (<.001)	0.349
Health care utilization in last 6 months					
Currently receive other addiction treatments (Yes)	207	89 (43)	78 (37.7)	1.639 (.20)	0.694
HIV in last 6 months					
HIV risky behavior (Yes)	207	76 (36.7)	65 (31.4)	1.818 (.18)	0.667
HIV testing (Yes)	206	81 (39.3)	116 (56.3)	33.03 (<.001)	—[b]

[a] Numbers lower than 1 mean reductions of the events from baseline to 6 months.

[b] The odds ratio for this variable cannot be calculated because patients' HIV testing status was considered current at 6 months if they had been tested at baseline; that is, there were zero patients considered tested at pretest and not tested at 6 months.

Figure 2. Exchange among Seva patients in the Bronx.



From: Janice
Date: 03/11/2016 4:41 AM **Title:** Ruff time Thursday
once again I wanted to pick up Crack not to make light of it the name alone speaks for its self. broken , garbage, destroyed . I want to be whole not garbage. I fill really confused some times. the word commitment comes to my mind . I not willing. the true hurts. thank for listening, I needed to share. I didn't pick up.

From: Dennis 03/11/2016 5:51 AM
Morning Janice, You did the right thing by sharing. I've been having a lot of difficulty staying focused these past 9 months, but I always share with the SEVA family. One thing I like about this is that I can be anonymous, but I'm also human with feelings & I'll always get feedback whether I agree with it or not. I hope this helped a little. Ciao.

From: Janice 03/11/2016 5:53 AM
thank you

From: Dennis 03/11/2016 6:03 AM
I figured it works for me, it help you when it gets tough.

From: Josie 03/11/2016 6:03 AM
you are on the right path Janice. one step at a time. confidence within you will help you conquer the urges. good luck and blessings

From: Dennis 03/11/2016 6:13 AM
That's right, stay focused. It's not easy, but if nothing else, you have the SEVA family, so please take full advantage of it as often as necessary. Ciao.

From: Mary 03/11/2016 6:25 AM
janice I'm glad you shared that with us, and happy you didn't pick up, our higher power was standing besides you, and he was not going to allow for you to pick up. try to.make a meeting today, that will work, God bless you. remember One Day At A Time.

From: Trent 03/11/2016 6:27 AM
You are not alone. This disease is a everyday battle. It's not kind, prejudice or unforgiving. It's objective is to kill, destroy and keep us in a state of confusion. It doesn't have a schedule to attack. It just comes for you without warning. Be prepared. Our weapons : meetings, literature, groups, prayer and contacts. I don't know how long SEVA has existed. To me it's inovated and something new. You used it, reached out to others and didn't pick up. I'd say you were prepared. Never give up or in. The disease of addiction doesn't. Have a nice day and enjoy life as it was meant to be.

A second unexpected development affected clinician adoption: We never succeeded in incorporating Seva data into the EHR at each clinic. The interoperability of EHRs is a widely recognized problem. In fact, the same system (eg, Epic Systems, developed by Epic Systems Corporation) may function very differently at one clinic versus another. The technical challenge in getting Seva data into the EHR of 3 health systems proved insurmountable in the context of this implementation research project. Instead, clinicians who wanted to review Seva data had to go to a website outside the EHR.

At each site, the leaders of Seva implementation and operation logged into Seva regularly during the 24 months of enrollment and implementation. At Site 1, 2 other staff clinicians besides the clinic Seva leader from behavioral health were most involved with Seva. The 3 clinicians at Site 1—one primary care physician and 2 behavioral health staff members—logged in an average of 0.6, 0.5, and 1.0 days per week. The Seva clinic leaders at Sites 2 and 3 were also the main users of Seva. The director of innovation at Site 2 logged in an average of 1.7 days per week, and a member of the behavioral health group at Site 3 logged in an average of 4.1 days per week. At two sites, behavioral health providers made most referrals of patients to Seva: 92% (89/97) of referrals at Site 1 and 76% (54/71) at Site 3. At Site 2, on the other hand, where we observed especially strong leadership and operational support for integrating behavioral with medical health, medical providers made 92% (92/100) of referrals (see Table 5).

Implementation

Table 5 also shows the percent of implementation goals and milestones completed. At the first two sites, we largely executed our implementation plan successfully in the preimplementation and implementation stages. At the third site, we struggled during preimplementation, primarily because of delays in getting the institutional review board’s (IRB) approval, which cost 6 additional months. Cost, which was specified in the protocol as another element of implementation, is reported below.

Figure 3. Percentage of patients who logged onto Seva at least once per study week. Patients were excluded from analysis at the point when they were removed from the study (eg, if they lost their phone, died, or were incarcerated).

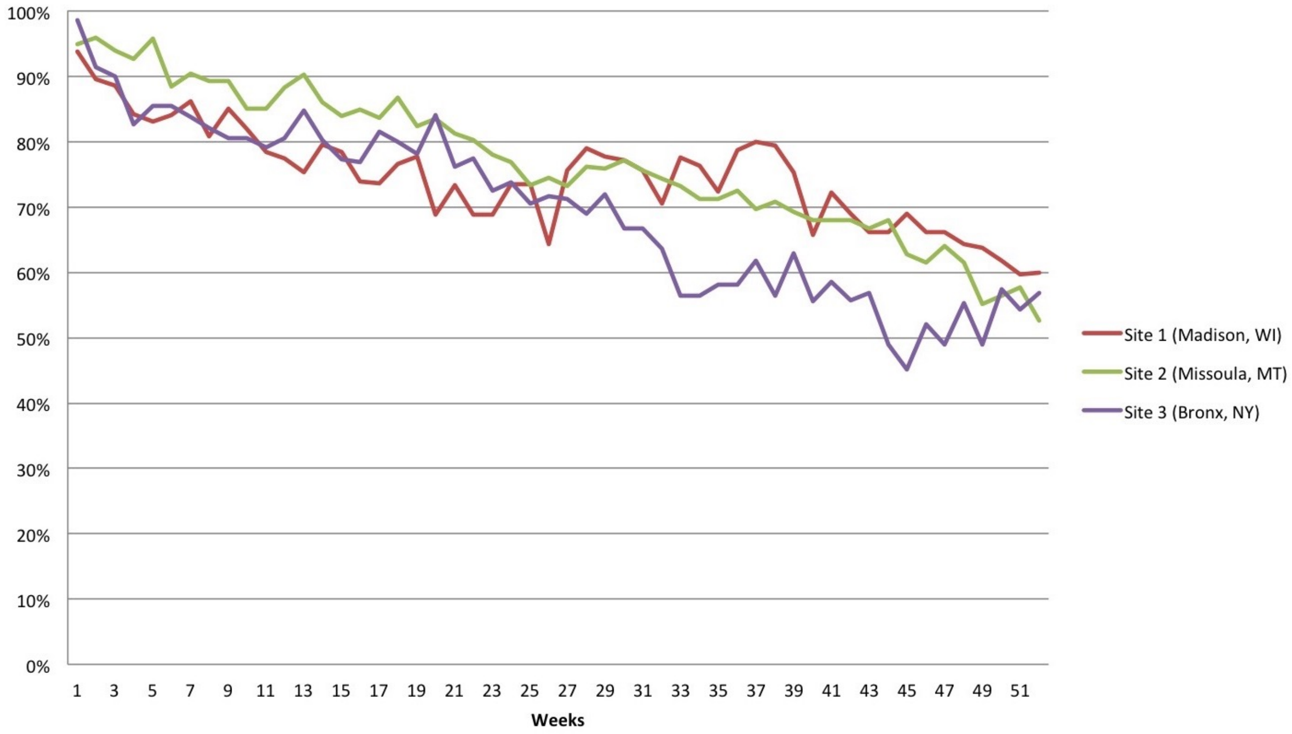


Table 5. Adoption and implementation outcomes.

Measures	Site 1 (Madison, WI)	Site 2 (Missoula, MT)	Site 3 (Bronx, NY)
Total number of Seva patients	97	100	71
Total number of primary care clinical staff at site[a]	70	74	27
Number of patients referred to Seva by primary care clinical staff out of total number of Seva patients, n (%)	8 (8)	92 (92)	17 (24)
Total number of behavioral health providers at site[b]	10	8	11
Number of patients referred to Seva by behavioral health providers out of total number of Seva patients, n (%)	89 (92)	8 (8)	54 (76)
Time to completion of project phases (months)			
Preimplementation (12 months planned)	12	12	18
Implementation (12 months planned)	18	13	16
Maintenance (12 months planned)	2	4	1
Number and percentage of implementation goals and milestones completed, n (%)			
Preimplementation (% of monthly implementation plan milestones reached)	12 (100)	12 (100)	12 (100)
Implementation (% of 100 patient enrollment goal at each site)	97 (97)	100 (100)	71 (71)
Maintenance (% of 12 monthly coaching follow up calls completed)	2 (17)	4 (33)	1 (8)

[a] Clinical staff members were physicians, residents, physician assistants, nurse practitioners, registered nurses, licensed practical nurses, and medical assistants.

[b] Behavioral health providers were licensed medical social workers, licensed mental health counselors, licensed clinical social workers, doctorate-level psychologists, and physicians.

Table 6. Cost analysis (US $).

Type of costs	Site 1 (Madison, WI)	Site 2 (Missoula, MT)	Site 3 (Bronx, NY)
Operating costs (patient and clinic)	113,636	117,150	83,177
Implementation costs (clinic)	9948	23,345	28,121
Total costs for clinic	123,584	140,495	111,298
Cost/patient	1274	1405	1568

Our implementation plans and their execution arose from baseline assessments of readiness for implementation, sustainability potential, and technology acceptance by clinic staff members who worked on the implementation plan, as well as the information that researchers gained from directly interacting with clinic staff. The plans for implementation involved using an organizational coach to help clinics through the 4 phases (initiate, prepare, improve, and implement), using an initial in-person visit and then monthly phone calls between the coach and the clinic change team. Although the organizational coach was heavily involved in the first two phases, another member of the research team who was most familiar with operating Seva became the primary contact for the clinics starting with the third phase. This researcher monitored the Seva discussion groups and gave hands-on, practical advice in response to questions from clinicians. Instead of having monthly coaching calls, clinician-researcher contact consisted mostly of short, frequent, ad hoc phone and email communications between the main clinician users and the research team's site coordinator in response to specific patient issues and technical questions. These were the types of problems clinics wanted help in solving, rather than helping make organizational change, which was the primary type of help offered by the coach. Thus, we adapted our implementation plan at each site to remove monthly coaching calls during the implementation phase and instead focused on the enrollment of patients as our primary implementation goal.

Maintenance

Maintenance, defined as continued use of Seva after the 12 months during which patient phones and data plans were paid for, was low at all 3 sites. Patients were allowed to keep their phones and could continue accessing Seva by using Wi-Fi or paying for service themselves with a new phone number. Although we did not track whether patients paid for new data plans independently, we did track Seva use after the 12 months of paid use. Use after 12 months declined gradually to zero once the last recruited patient reached the end of the 12-month period of paid phone use. The decline in use limited our ability to collect follow-up patient surveys and led us to abandon our attempts to collect the phone survey we planned to collect 6 months after each patient's 12-month intervention period to gauge maintenance. The decline in clinician use of Seva mirrored patient decline. That is, as fewer patients used the system, or if only a few later-enrolled patients were using it, clinicians logged in less because Seva had little patient activity for them to see.

Sites did not continue to use Seva after the study for at least two reasons. First, we could not resolve issues related to transitioning from a research study to ongoing use of an mHealth system. In particular, establishing procedures for consenting patients who wanted to use their own phones to access Seva outside the research protocol, and commingling these patients with patients from the research study, proved challenging. Second, the National Institute of Health grant funding ended, and none of the 3 clinics made arrangements to pay for mobile phones and data plans afterwards. Offering Seva only to patients who can cover their own mobile phone costs could have been a condition of eligibility, but this choice would have shifted the cost to patients and restricted access for low-income patients. It may also have reduced patients' motivation to use the system, because patients reported that receiving a mobile phone was a strong incentive to use Seva.

Without ongoing funding for patients to use the system, use of Seva declined significantly and patients ultimately became unreachable to the research staff (note that patient surveys were administered over the phone, using the number associated with the patient's mobile phone). This led us to abandon our attempts to collect the phone survey we planned to collect 6 months after each patient's 12-month intervention period to gauge maintenance.

Cost

This analysis addresses cost from the perspective of an FQHC clinic administrator—a comprehensive economic analysis of the study will be reported separately. This analysis includes operating costs (eg, mobile phones, data plans, clinician time for monitoring Seva use, information technology [IT] staff time) and the costs of executing the implementation strategy (eg, site visits, coaching calls). Costs were tracked for 36 months across each project phase (preimplementation, implementation, and maintenance). Costs are broken down as follows: system operating costs; implementation costs per clinic; and overall cost per patient and per clinic (see Table 6). All costs are given in US dollars. Operating costs per patient were estimated at US $1185, which covers US $200 for a mobile phone; US $720 for a voice and data plan (US $60/month x 12 months); US $135 in clinical staff time for patient identification, recruitment, and training (based on observation and interviews with staff—this 2-stage process took an average of 1.5 hours per patient x US $90/hour); and US $130 in staff time for monitoring patients' use of Seva (based on server logs, clinical and research staff spent 0.12 hours per patient per month monitoring patients; 0.12 hours per patient x 12 months x US $90/hour). We conducted interviews with IT staff members to derive estimates of system operating costs of US $8,000/per clinic over 36 months, which covers costs such as technical support for users, server hosting, and software updates to the system. Implementation costs were estimated at US $10,350 per clinic, which includes US $8,100 for coaching time and expenses associated with site visits (3 visits per clinic at US $500/day, including travel costs of US $1200 per visit) and US $2250 for monthly follow up via email and phone (36 hours x US $62.50/hour).

Total cost per clinic averaged approximately US $124,000 across the 3 clinics; cost per patient averaged approximately US $1,400. For comparison, the average cost for an episode of outpatient addiction treatment among 21 addiction treatment programs surveyed in 2008 (the year with the most recent available cost data) was US $2325 [37].

Discussion

Principal Findings

To summarize our findings related to the three research questions that were the focus of the study:

> RQ1: How can Seva be implemented in primary care settings efficiently and effectively?

The study offers the following 4 lessons about implementing an mHealth system in primary care:

First, plan and have a budget for working extensively with clinic IT staff to integrate mHealth data into the EHR. This very challenging task is essential for integrating mHealth into primary care because it makes the mHealth data part of rather than separate from the data clinicians expect to see as they treat patients. Second, work with clinic staff to figure out how the mHealth system will fit into the clinic's existing workflow. All 3 clinics in this study chose to appoint one clinician to monitor Seva data and alert fellow clinicians as needed about important changes in a patient's recovery. Wider and deeper integration would result from each clinician routinely monitoring data from the mHealth system for his or her patients, just as he or she monitors data related to other chronic conditions such as diabetes. This level of involvement may be unrealistic, however, given the time pressures on primary care staff, and assigning routine monitoring to other clinical staff has been done effectively in other studies we have conducted [38,39]. Third, ensure that the questions clinical staff members have as they operate the mHealth system can get rapid responses from the mHealth developers. Most questions clinicians had in this study were day-to-day operational issues (eg, how to enter a new patient into the system) that were ideally addressed in the moment they occurred by quickly calling the research team. Fourth, to assure sustained use, address cost. Two of the 3 clinics wanted to maintain using Seva based on patient and clinician feedback and, at one site, cost savings from reduced hospital admissions and ER visits. This last clinic used the cost savings to make a case to an insurer to pay for Seva, a process still under way at this writing.

> RQ2: To what extent do patients and staff accept and use Seva?

Patient use was exceptionally high compared with continued use of most mHealth apps, although patient use declined steeply after funding for the phones and data plans ended. Clinician use was low compared with patient use because, as stated above, clinicians worried about being responsible for data from Seva and they had to view Seva data in a separate website rather than in the EHR. Of the two aspects of integration we examined—the integration of behavioral health into primary care, and the integration of mHealth into the treatment of addiction in primary care—the second was more successful than the first. Although the treatment of patients suffering from addiction and the use of Seva remained mainly the province of behavioral health providers in this study, the integration of mHealth into addiction treatment was successful if judged by the high levels of patient use.

> RQ3: How does Seva affect clinical care for patients and staff?

This study showed the potential of patient peer support in encouraging treatment adherence. Patient peer support is unusual in primary care. It is also a type of care that does not add to, and may reduce, clinician burden. Clinicians who used Seva were generally enthusiastic about it, as demonstrated by 2 clinics wanting to continue using the system after grant funding ended, even though these intentions have thus far gone unfulfilled. The number of clinicians involved, though, is too small to warrant generalizations.

Comparisons With Prior Work

This study—which involved 3 unaffiliated primary care clinics enrolling nearly 300 patients in the same mHealth system—is the most comprehensive implementation research trial focused on the use of mHealth in primary care yet conducted in the US health care system. Implementation research on mHealth has been focused almost entirely on developing countries in Africa, where mHealth is usually used as a replacement for standard health care and operates independently of any health care system [40].

Prior studies focused on providing addiction treatment to primary care patients have also been rare; we found only 2 clinical trials for primary care SUD interventions [41,42]. Neither of these trials used mHealth. Both addressed illicit drug use and prescription drug abuse rather than AUD and neither had an effect compared with the control group.

We observed significant reductions in drinking and substance use among patients using Seva. These are promising reductions compared with other outpatient AUD treatments. For example, Project MATCH showed that psychosocial treatments for alcoholism are not particularly effective [43], although more recent studies of cognitive-behavioral therapy for SUDs have demonstrated efficacy [44]. A Cochrane review of the effectiveness of naltrexone, a commonly administered medication for AUD, found a decreased risk of heavy drinking by 17% compared with the placebo group and a decrease in the number of drinking days of about 4% [45].

Granted, this study was an implementation research trial not specifically designed to retest the effectiveness of Seva. A rigorously designed randomized controlled trial would be required to definitively demonstrate the effectiveness of Seva within a substance-using primary care population.

Limitations

The study has some limitations. (1) The primary outcome, risky drinking days, is a self-reported measure. (2) Selection bias could have affected which patients participated in the study. Clinicians were advised to enroll any eligible patients they thought might benefit from Seva, based on their clinical judgment. It was not possible for us to tell if this led clinicians to favor some patients over others (eg, those with more education). (3) We could not retrieve racial and ethnic data for SUD patients who were eligible to participate in the study but were not enrolled, and we did not collect data about education and socioeconomic status on Seva patient surveys, limiting our understanding of the representativeness of the patient sample. (4) We lost an opportunity to learn about sustainability from the patient’s perspective because we could not reach patients for the survey we planned to collect 6 months after each patient’s 12-month intervention period. (5) The study would be hard to replicate because it reports on a mobile phone app that, like almost all such apps, changes over time rather than remaining stable. For example, since the start of this study, the app has been rebuilt to update it to current accessibility and security standards, add features (My Motivation), improve features (the automatic tailoring of feedback in response to weekly surveys), and add content for opioid addiction.

Challenges

The current findings suggest that mHealth faces several formidable challenges to widespread implementation. The first is integrating mHealth with EHRs (eg, Epic Systems), described above. Clinicians in this study had to view Seva data by going to a website outside the EHR, requiring already overburdened clinicians to make extra mouse-clicks. In addition, logistical and patient privacy concerns limited our ability to use EHR data for evaluation purposes. For example, we could not obtain participants’ attendance at primary care visits, which we planned to include in the health care utilization analyses. The poor interoperability and accessibility of EHRs are likely to generalize to other researchers and mHealth developers, making research challenging to conduct and limiting the usefulness of mHealth systems for clinicians.

Another key challenge relates to the difficulty of enrolling patients with SUDs. Clinics reported that it often took multiple follow-up calls to get patients into the clinic to get informed consent and conduct training, even though patients were often excited about getting a mobile phone. The effort required to enroll patients might have dampened enthusiasm for enrollment beyond the first 100 patients that the budget provided for.

FQHCs are required by law to provide behavioral health services. We do not know how Seva might function in a primary care clinic without designated behavioral health staff. Although Seva was somewhat integrated into behavioral health, it did not deeply penetrate physicians' treatment of patients, as reported previously about Site 1 [26] and observed in Sites 2 and 3, perhaps reflecting the deep divides between primary care and behavioral health mentioned in the Introduction [15].

Finally, our experience has revealed several fundamental questions about the role of mHealth in primary care: Is the value of patient peer support for behavioral health sufficient to make the costs of embedding an mHealth system such as Seva into a clinic’s operations worthwhile? Might it instead suffice for patients to use mHealth systems on their own, based upon the recommendation of primary care clinicians? If integrating mHealth into clinic systems is deemed worthwhile, who bears the costs? Certain costs, such as those for mobile phones and data plans, could be borne by patients, some of whom already pay for phones and data plans out of pocket. In this study, patient use of Seva and survey follow-up rates dropped significantly when research funding stopped paying for data plans. Yet, these costs account for only about half the total cost of the system, and lower-income patients often have pay-as-you-go data plans that may not work with the data requirements of an mHealth system such as Seva. Volunteer peer mentors could potentially act as monitors, as clinicians and researchers did in this project, thereby reducing the cost to clinics. Indeed, similar volunteer roles are essential to the Alcoholics Anonymous model (eg, “sponsors”).

Conclusions

mHealth has the potential to transform health care, and given the enormous cost of health care, we need to make effective use of every available resource. In contrast to the seemingly inexorable rise in health care costs over time, the cost of technology tends to decrease in accordance with Moore’s Law, which posits that computing capability roughly doubles every 2 years. Our experience illustrates that mHealth can engage patients suffering from addiction in ways that benefit patients without adding substantial burden on health care providers. Although challenges remain, thoughtful deployment of mHealth could improve the treatment of addiction in primary care and might also improve the treatment of other chronic conditions that have prominent behavioral components (eg, diabetes). In so doing, mHealth could transcend the physically local and professionally controlled systems that characterize the US health care system.

Acknowledgments

This research was funded by the National Institute on Drug Abuse grant number R01-DA034279. The funder had no role in any aspect of the development, conduct, analysis, or reporting of the study. The authors wish to thank the dedicated clinicians and staff members at the 3 clinics where the study was conducted, in particular, Chantelle Thomas from Access Community Health Center in Madison, WI; Virna Little, Victoria Ward, and Tom McCarry from the Institute for Family Health, Center for Counseling, at Walton, Bronx, NY; and Mary Jane Nealon from Partnership Health Center in Missoula, MT. The project could not have succeeded without their dedication and support. We also thank Judith Ganch, Joseph Glass, Andrew Isham, Adam Maus, Helene McDowell, Klaren Pe-Romashko, and Victoria Ward for their important contributions to the project.

Conflicts of Interest

Authors DG, FM, RJ, and AQ have a shareholder interest in CHESS Mobile Health, a public benefit corporation that disseminates technology to the specialty addiction treatment system. CHESS Mobile Health did not develop Seva. The relationship between the authors and CHESS Mobile Health is managed by the authors and the University of Wisconsin–Madison's Conflict of Interest Committee. LM is affiliated with HealthSim, LLC—a small business that developed the Web-based Therapeutic Education System (TES). The relationship is extensively managed by LM and her academic institution. All other authors declare that they have no competing interests.

References

1. Bauer AM, Rue T, Keppel GA, Cole AM, Baldwin L, Katon W. Use of mobile health (mHealth) tools by primary care patients in the WWAMI region Practice and Research Network (WPRN). J Am Board Fam Med 2014;27(6):780-788 [FREE Full text] [doi: 10.3122/jabfm.2014.06.140108] [Medline: 25381075]
2. Singh K, Drouin K, Newmark LP, Lee J, Faxvaag A, Rozenblum R, et al. Many mobile health apps target high-need, high-cost populations, but gaps remain. Health Aff (Millwood) 2016 Dec 01;35(12):2310-2318. [doi: 10.1377/hlthaff.2016.0578] [Medline: 27920321]
3. Wicks P, Chiauzzi E. 'Trust but verify'—five approaches to ensure safe medical apps. BMC Med 2015;13:205 [FREE Full text] [doi: 10.1186/s12916-015-0451-z] [Medline: 26404791]
4. Ades J. Mpulsemobile. Healthcare apps: the good, the bad, the ugly URL: https://mpulsemobile.com/2016/07/good-bad-ugly-healthcare-apps/ [WebCite Cache ID 6t6bOSphh]
5. Torous J, Chan S, Luo J. Psychnews.psychiatryonline. 2016. Are your patients using 'digital supplements'? URL: http://psychnews.psychiatryonline.org/doi/full/10.1176%2Fappi.pn.2016.11b16 [WebCite Cache ID 6t6c6W2Kp]
6. 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]
7. 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]
8. Steinhubl SR, Muse ED, Topol EJ. Can mobile health technologies transform health care? J Am Med Assoc 2013 Dec 11;310(22):2395-2396. [doi: 10.1001/jama.2013.281078] [Medline: 24158428]
9. Quanbeck A, Chih MY, Isham A, Gustafson D. Mobile delivery of treatment for alcohol use disorders: a review of the literature. Alcohol Res 2014;36(1):111-122 [FREE Full text] [Medline: 26259005]
10. Riper H, Spek V, Boon B, Conijn B, Kramer J, Martin-Abello K, et al. Effectiveness of E-self-help interventions for curbing adult problem drinking: a meta-analysis. J Med Internet Res 2011;13(2):e42 [FREE Full text] [doi: 10.2196/jmir.1691] [Medline: 21719411]
11. Campbell W, Hester RK, Lenberg KL, Delaney HD. Overcoming Addictions, a web-based application, and SMART Recovery, an online and in-person mutual help group for Problem drinkers, part 2: six-month outcomes of a randomized controlled trial and qualitative feedback from participants. J Med Internet Res 2016 Oct 04;18(10):e262 [FREE Full text] [doi: 10.2196/jmir.5508] [Medline: 27701064]
12. Berrouiguet S, Baca-García E, Brandt S, Walter M, Courtet P. Fundamentals for future mobile-health (mHealth): a systematic review of mobile phone and web-Based text messaging in mental health. J Med Internet Res 2016 Jun 10;18(6):e135 [FREE Full text] [doi: 10.2196/jmir.5066] [Medline: 27287668]
13. World Health Organization. E-health technologies and substance abuse URL: http://www.who.int/substance_abuse/activities/ehealth/en/ [WebCite Cache ID 6vHzMRZzq]
14. Klein S, Hostetter M. The Commonwealth Fund. 2014. In focus: integrating behavioral health and primary care URL: http://www.commonwealthfund.org/publications/newsletters/quality-matters/2014/august-september/in-focus [WebCite Cache ID 6t6ceiR8u]
15. Awoyinka L, Gustafson D, Johnson R. Using technology to integrate behavioral health into primary care. In: Marsch L , Lord S , Dallery J , editors. Behavioral Healthcare and Technology: Using Science-Based Innovations to Transform Practice. New York: Oxford University Press; 2014.
16. Dennis M, Scott CK. Managing addiction as a chronic condition. Addict Sci Clin Pract 2007 Dec;4(1):45-55 [FREE Full text] [Medline: 18292710]
17. Pating DR, Miller MM, Goplerud E, Martin J, Ziedonis DM. New systems of care for substance use disorders: treatment, finance, and technology under health care reform. Psychiatr Clin North Am 2012 Jun;35(2):327-356. [doi: 10.1016/j.psc.2012.03.004] [Medline: 22640759]
18. Capoccia VA, Grazier KL, Toal C, Ford 2nd JH, Gustafson DH. Massachusetts's experience suggests coverage alone is insufficient to increase addiction disorders treatment. Health Aff (Millwood) 2012 May;31(5):1000-1008 [FREE Full text] [doi: 10.1377/hlthaff.2011.0326] [Medline: 22566439]
19. DeRubeis RJ, Crits-Christoph P. Empirically supported individual and group psychological treatments for adult mental disorders. J Consult Clin Psychol 1998 Feb;66(1):37-52. [Medline: 9489261]

20. Molfenter T. Reducing appointment no-shows: going from theory to practice. Subst Use Misuse 2013 Jun;48(9):743-749 [FREE Full text] [doi: 10.3109/10826084.2013.787098] [Medline: 23607670]
21. 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. J Am Med Assoc Psychiatry 2014 May;71(5):566-572 [FREE Full text] [doi: 10.1001/jamapsychiatry.2013.4642] [Medline: 24671165]
22. Bickel WK, Marsch LA, Buchhalter AR, Badger GJ. Computerized behavior therapy for opioid-dependent outpatients: a randomized controlled trial. Exp Clin Psychopharmacol 2008 Apr;16(2):132-143 [FREE Full text] [doi: 10.1037/1064-1297.16.2.132] [Medline: 18489017]
23. Glass JE, McKay JR, Gustafson DH, Kornfield R, Rathouz PJ, McTavish FM, et al. Treatment seeking as a mechanism of change in a randomized controlled trial of a mobile health intervention to support recovery from alcohol use disorders. J Subst Abuse Treat 2017 Jun;77:57-66. [doi: 10.1016/j.jsat.2017.03.011] [Medline: 28476273]
24. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care 2012 Mar;50(3):217-226 [FREE Full text] [doi: 10.1097/MLR.0b013e3182408812] [Medline: 22310560]
25. Quanbeck AR, Gustafson DH, Marsch LA, McTavish F, Brown RT, Mares M, et al. Integrating addiction treatment into primary care using mobile health technology: protocol for an implementation research study. Implement Sci 2014 May 29;9:65 [FREE Full text] [doi: 10.1186/1748-5908-9-65] [Medline: 24884976]
26. Mares M, Gustafson DH, Glass JE, Quanbeck A, McDowell H, McTavish F, et al. Implementing an mHealth system for substance use disorders in primary care: a mixed methods study of clinicians' initial expectations and first year experiences. BMC Med Inform Decis Mak 2016 Sep 29;16(1):126 [FREE Full text] [doi: 10.1186/s12911-016-0365-5] [Medline: 27687632]
27. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [Medline: 10474547]
28. Wen K, Gustafson DH, Hawkins RP, Brennan PF, Dinauer S, Johnson PR, et al. Developing and validating a model to predict the success of an IHCS implementation: the Readiness for Implementation Model. J Am Med Inform Assoc 2010;17(6):707-713 [FREE Full text] [doi: 10.1136/jamia.2010.005546] [Medline: 20962135]
29. Maher L, Gustafson D, Evans A. Webarchive.nationalarchives.gov.uk. Leicester, England: British National Health Service Modernization Agency; 2004. Sustainability, model and guide URL: http://webarchive.nationalarchives.gov.uk/20160805122935/http://www.nhsiq.nhs.uk/media/2757778/nhs_sustainability_model_-_february_2010_1_.pdf [accessed 2017-08-30] [WebCite Cache ID 6t6wxXbsT]
30. Legris P, Ingham J, Collerette P. Why do people use information technology? A critical review of the technology acceptance model. Inf Manage 2003 Jan;40(3):191-204. [doi: 10.1016/S0378-7206(01)00143-4]
31. Cranford JA, McCabe SE, Boyd CJ. A new measure of binge drinking: prevalence and correlates in a probability sample of undergraduates. Alcohol Clin Exp Res 2006 Nov;30(11):1896-1905 [FREE Full text] [doi: 10.1111/j.1530-0277.2006.00234.x] [Medline: 17067355]
32. Hays RD, Bjorner JB, Revicki DA, Spritzer KL, Cella D. Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. Qual Life Res 2009 Sep;18(7):873-880 [FREE Full text] [doi: 10.1007/s11136-009-9496-9] [Medline: 19543809]
33. Chamberlain P, Brown CH, Saldana L. Observational measure of implementation progress in community based settings: the Stages of Implementation Completion (SIC). Implement Sci 2011 Oct 06;6:116 [FREE Full text] [doi: 10.1186/1748-5908-6-116] [Medline: 21974914]
34. Center for Behavioral Health Statistics and Quality. Substance Abuse and Mental Health Services Administration. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2016. 2015 National Survey on Drug Use and Health: Detailed Tables URL: https://www.samhsa.gov/data/sites/default/files/NSDUH-DetTabs-2015/NSDUH-DetTabs-2015/NSDUH-DetTabs-2015.pdf [WebCite Cache ID 6wi6oMKUd]
35. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
36. Baldwin JL, Singh H, Sittig DF, Giardina TD. Patient portals and health apps: pitfalls, promises, and what one might learn from the other. Healthc (Amst) 2016 Oct 3;5(3):81-85 [FREE Full text] [doi: 10.1016/j.hjdsi.2016.08.004] [Medline: 27720139]
37. French MT, Popovici I, Tapsell L. The economic costs of substance abuse treatment: updated estimates and cost bands for program assessment and reimbursement. J Subst Abuse Treat 2008 Dec;35(4):462-469 [FREE Full text] [doi: 10.1016/j.jsat.2007.12.008] [Medline: 18294803]
38. DuBenske LL, Gustafson DH, Shaw BR, Cleary JF. Web-based cancer communication and decision making systems: connecting patients, caregivers, and clinicians for improved health outcomes. Med Decis Making 2010;30(6):732-744 [FREE Full text] [doi: 10.1177/0272989X10386382] [Medline: 21041539]
39. Dubenske LL, Chih M, Dinauer S, Gustafson DH, Cleary JF. Development and implementation of a clinician reporting system for advanced stage cancer: initial lessons learned. J Am Med Inform Assoc 2008;15(5):679-686 [FREE Full text] [doi: 10.1197/jamia.M2532] [Medline: 18579837]

40. Aranda-Jan CB, Mohutsiwa-Dibe N, Loukanova S. Systematic review on what works, what does not work and why of implementation of mobile health (mHealth) projects in Africa. BMC Public Health 2014;14:188 [FREE Full text] [doi: 10.1186/1471-2458-14-188] [Medline: 24555733]
41. Saitz R, Palfai TP, Cheng DM, Alford DP, Bernstein JA, Lloyd-Travaglini CA, et al. Screening and brief intervention for drug use in primary care: the ASPIRE randomized clinical trial. J Am Med Assoc 2014 Aug 06;312(5):502-513 [FREE Full text] [doi: 10.1001/jama.2014.7862] [Medline: 25096690]
42. Roy-Byrne P, Bumgardner K, Krupski A, Dunn C, Ries R, Donovan D, et al. Brief intervention for problem drug use in safety-net primary care settings: a randomized clinical trial. J Am Med Assoc 2014 Aug 06;312(5):492-501 [FREE Full text] [doi: 10.1001/jama.2014.7860] [Medline: 25096689]
43. Cutler RB, Fishbain DA. Are alcoholism treatments effective? The Project MATCH data. BMC Public Health 2005 Jul 14;5:75 [FREE Full text] [doi: 10.1186/1471-2458-5-75] [Medline: 16018798]
44. McHugh RK, Hearon BA, Otto MW. Cognitive behavioral therapy for substance use disorders. Psychiatr Clin North Am 2010 Sep;33(3):511-525 [FREE Full text] [doi: 10.1016/j.psc.2010.04.012] [Medline: 20599130]
45. Rösner S, Hackl-Herrwerth A, Leucht S, Vecchi S, Srisurapanont M, Soyka M. Opioid antagonists for alcohol dependence. Cochrane Database Syst Rev 2010 Dec 08(12):CD001867. [doi: 10.1002/14651858.CD001867.pub2] [Medline: 21154349]

Abbreviations

EHR: electronic health record
ER: emergency room
FQHC: federally qualified health center
IT: information technology
mHealth: mobile health
QoL: quality of life
RE-AIM: Reach, Effectiveness, Adoption, Implementation, and Maintenance

Edited by G Eysenbach; submitted 12.09.17; peer-reviewed by P Wicks, K Claborn; comments to author 05.10.17; revised version received 30.11.17; accepted 22.12.17; published 30.01.18.

Please cite as:
Quanbeck A, Gustafson DH, Marsch LA, Chih MY, Kornfield R, McTavish F, Johnson R, Brown RT, Mares ML, Shah DV
Implementing a Mobile Health System to Integrate the Treatment of Addiction Into Primary Care: A Hybrid Implementation-Effectiveness Study
J Med Internet Res 2018;20(1):e37
URL: http://www.jmir.org/2018/1/e37/
doi:10.2196/jmir.8928
PMID:29382624

Original Paper

Web-Delivered Cognitive Behavioral Therapy for Distressed Cancer Patients: Randomized Controlled Trial

Suzanne K Chambers[1,2,3,4,5], PhD; Lee M Ritterband[6,7], PhD; Frances Thorndike[7], PhD; Lisa Nielsen[2], MPH; Joanne F Aitken[1,2,8], PhD; Samantha Clutton[2], M Psych (Clin); Paul A Scuffham[1], PhD; Philippa Youl[9,10], PhD; Bronwyn Morris[11], PhD; Peter D Baade[1,2,12], PhD; Jeff Dunn[1,2,8,13], PhD

[1]Menzies Health Institute Queensland, Griffith University, Gold Coast, Australia
[2]Cancer Council Queensland, Brisbane, Australia
[3]Prostate Cancer Foundation of Australia, Sydney, Australia
[4]Health and Wellness Institute, Edith Cowan University, Perth, Australia
[5]Centre for Clinical Research, The University of Queensland, Brisbane, Australia
[6]University of Virginia, Charlottesville, VA, United States
[7]BeHealth Solutions, Charlottesville, VA, United States
[8]Institute for Resilient Regions, University of Southern Queensland, Springfield, Australia
[9]University of Sunshine Coast, Sippy Downs, Australia
[10]School of Public Health and Social Work, Queensland University of Technology, Brisbane, Australia
[11]Menzies School of Health Research, Darwin, Australia
[12]School of Mathematical Sciences, Queensland University of Technology, Brisbane, Australia
[13]School of Social Science, The University of Queensland, Brisbane, Australia

Corresponding Author:
Suzanne K Chambers, PhD
Menzies Health Institute Queensland
Griffith University
Griffith Health Centre
Gold Coast,
Australia
Phone: 61 7 5678 8664
Email: suzanne.chambers@griffith.edu.au

Abstract

Background: Web-based interventions present a potentially cost-effective approach to supporting self-management for cancer patients; however, further evidence for acceptability and effectiveness is needed.

Objective: The goal of our research was to assess the effectiveness of an individualized Web-based cognitive behavioral therapy (CBT) intervention on improving psychological and quality of life outcomes in cancer patients with elevated psychological distress.

Methods: A total of 163 distressed cancer patients (111 female, 68.1%) were recruited through the Queensland Cancer Registry and the Cancer Council Queensland Cancer Helpline and randomly assigned to either a Web-based tailored CBT intervention (CancerCope) (79/163) or a static patient education website (84/163). At baseline and 8-week follow-up we assessed primary outcomes of psychological and cancer-specific distress and unmet psychological supportive care needs and secondary outcomes of positive adjustment and quality of life.

Results: Intention-to-treat analyses showed no evidence of a statistically significant intervention effect on primary or secondary outcomes. However, per-protocol analyses found a greater decrease for the CancerCope group in psychological distress ($P=.04$), cancer-specific distress ($P=.02$), and unmet psychological care needs ($P=.03$) from baseline to 8 weeks compared with the patient education group. Younger patients were more likely to complete the CancerCope intervention.

Conclusions: This online CBT intervention was associated with greater decreases in distress for those patients who more closely adhered to the program. Given the low costs and high accessibility of this intervention approach, even if only effective for subgroups of patients, the potential impact may be substantial.

Trial Registration: Australian New Zealand Clinical Trials Registry ACTRN12613001026718; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=364768&isReview=true (Archived by WebCite at http://www.webcitation.org/6uPvpcovl)

(J Med Internet Res 2018;20(1):e42) doi:10.2196/jmir.8850

KEYWORDS

cancer; mental health; psychological distress; randomized controlled trial (RCT); health services delivery

Introduction

In 2012 it was estimated that there were 14.1 million new cases of cancer diagnosed globally [1]. Estimates suggest that in 2030 this number will reach 21.6 million [1], a substantial increase in the cancer burden that will in turn increase demands on the health care system. In this regard, people affected by cancer frequently report heightened psychological distress [2] that for some persists for many years [3-5]. It is now well acknowledged that psychosocial care is an essential component of quality cancer care [6]. However, how to deliver evidence-based psychosocial care on a population basis, given the current and future predicted prevalence of cancer and increasingly limited health care resources, remains a challenge.

Approaches to more effectively deliver evidence-based psychosocial care include a low-intensity framework through which cost-effective services can be delivered. Low-intensity care models have as their guiding values the principles of equity and access, with tailoring to the extent and depth of need and use of flexible delivery methods [7]. Within this framework, self-management has been proposed as an effective method by which patient needs can be met [8,9]. Web-based interventions present a specific appeal here as a remotely delivered low-cost approach to supporting self-management with potential for widespread dissemination [10]. Indeed, Web-based programs have been found to be effective in promoting behavior change with regard to stress management, exercise, nutrition, and participation in health care [11,12]. There are, however, questions still to be answered about the acceptability and effectiveness of Web-based interventions to improve psychological outcomes for cancer patients.

Accordingly, we undertook a randomized controlled trial to assess the effectiveness of an individualized Web-based cognitive behavioral intervention (CancerCope) in cancer patients who have or are at risk of having elevated psychological distress. CancerCope was compared with a static patient education website with participants assessed over a 2-month period. We hypothesized that, relative to participants receiving patient education, participants receiving CancerCope would have lower psychological and cancer-specific distress, lower unmet psychological supportive care needs, higher positive adjustment, and improved quality of life.

Methods

Participants

Participants were recruited through the Queensland Cancer Registry (QCR), a population-based register of cancer diagnoses in Queensland, and the Cancer Council Queensland Cancer Helpline, a telephone information and support service. Eligible participants were adults who had been diagnosed with cancer who scored ≥4 on the Distress Thermometer [13] (indicating high distress or risk of high distress); were able to read and speak English; had no history of head injury, dementia or psychiatric illness; had no concurrent cancer; and had phone and Internet access. Participants recruited through the QCR had 2 additional eligibility criteria: consent from their diagnosing clinician to participate and having been diagnosed with melanoma or colorectal cancer within the last 6 months.

Intervention

Participants in the intervention arm were provided access to the CancerCope program, an online support program based on a 5-session telephone-based cognitive behavioral therapy intervention [8,14] and modified to include 6 cores covering: the cancer journey, understanding stress, managing worry, tackling problems, taking care (improving well-being), and moving forward. The cores consisted of educational information and expert videos from psychologists as well as stories and videos of 4 fictional characters on their cancer journey as a way to illustrate the different experiences of others. The program had high levels of interactivity to increase user engagement and systems to encourage use and self-management including personalized email reminders and feedback. Content was tailored in response to the participant’s needs as determined by their input, including assigned behavioral homework supported by the interactive components of the website. For example, users received tailored feedback based on distress scores and concerns. Users were also able to set personal goals and receive recommended goals. These were then tracked throughout the use of the program and could be modified by the user as needed.

Components that targeted challenges associated with cancer treatments (eg, pain, sleep disturbance, fatigue) were additionally selected if relevant. Cores were completed weekly over a 6-week period rather than available all at once, with ongoing access to the program provided for 12 months. Cores were marked as completed if the participant manually submitted them as complete. Screenshots of the CancerCope program can be found in Multimedia Appendix 1. A more detailed description of the program has been published elsewhere [15].

The control condition was a static patient education website containing information covering stress management skills, problem-solving approaches to cancer-related concerns, and patient education about a healthy lifestyle to promote wellness and optimize quality of life.

Participants were provided with the URL for the study website and a unique username and password that gave them individualized access to the program. Only the research team (project manager and staff involved with recruitment and follow-up) had access to participant information (including name and contact details) through a secure password-protected database. Data collected through online questionnaires were downloaded and saved on a secure password-protected server.

Study Integrity

Ethical approval was obtained from the Griffith University Human Research Ethics Committee (PSY/70/13/HREC) and Metro South Human Research Ethics Committee (HREC/13/QPAH/601). The study was guided by the Consolidated Standards of Reporting Trials (CONSORT) statement [16]. Randomization followed baseline assessment and occurred in blocks of 10, with each condition randomly generated 5 times within each block to ensure an unpredictable allocation sequence with equal numbers of participants in each group at the completion of each block. This sequence was undertaken by the project manager and concealed from investigators. Assessments were through self-report questionnaires. Primary analyses were intention to treat.

Materials

Baseline assessment was conducted by telephone. Follow-up assessment occurred after the intervention period (8 weeks) via online questionnaires accessed through the Web program.

Outcome Measures

Primary outcome measures included the Brief Symptom Inventory 18 [17], the Impact of Event Scale [18,19], and the Supportive Care Needs Survey Short Form 34 [20]. Higher scores on the first 2 measures indicated greater psychological or cancer-specific distress, respectively. Secondary outcome measures were the Posttraumatic Growth Inventory [21] and the Assessment of Quality of Life 8D [22]. Higher scores indicated greater benefit finding or quality of life, respectively. Process measures, as detailed next, were also included for the intervention arm.

Process Measures

Participants in the CancerCope condition completed 3 process measures following the 8-week intervention period. The Internet Evaluation and Utility Questionnaire assesses patients' experiences and perceptions of an Internet intervention [23,24]. The constructs measure ease of use, convenience, engagement, enjoyment, layout, privacy, satisfaction, acceptability, and perceptions of the Web program material in terms of usefulness, comprehension, credibility, likelihood of returning, mode of delivery, and helpfulness. Higher scores indicate more positive experiences and perceptions of the Web program. The Internet Intervention Adherence Questionnaire identifies obstacles and barriers that interfere with using Internet intervention programs [23,25]. Higher scores indicate the participant experienced more problems with the Web program. The Internet Impact and Effectiveness Questionnaire assesses patients' perceptions of the Internet intervention in terms of the program's effectiveness in resolving their targeted health condition. Perceived impact is measured in terms of helpfulness, knowledge gains, treatment effectiveness for self, treatment effectiveness for others, long-term effectiveness, quality of life, mood, physical activity, family relationships, peer relationships, social activity, school/work attendance, school/work performance, treatment implementation, goal orientation, confidence in ability to manage the health condition, relapse prevention, and service reduction [23,24]. Higher scores indicate greater impact and effectiveness.

Statistical Analyses

The study design involved a multivariate, 2-condition randomized controlled trial with repeated measures across time. A hierarchical linear model analysis was used to reflect this design in which measurement occasions (level 1) were nested within persons (level 2) and program differences were represented as a fixed effect at level 2 and the interaction with time suggested differential adjustment and distress trajectories for the 2 groups. The analysis examined the effect of study group (CancerCope and patient education) and time point (baseline and 2 months) on the specific primary and secondary outcome scores, including an interaction term between the 2 variables (study group and time point). We assessed differences in baseline demographic characteristics and baseline measures between respondents who did and did not complete the second questionnaire by performing multivariate backwards stepwise logistic regression analysis.

To facilitate an intention-to-treat analysis, multiple imputation (using 50 imputations) was used to impute missing data for those respondents who completed the baseline but not the follow-up assessment. The multiple imputation process involved regression of the relevant outcome variable with all the nonmissing values of the baseline outcome measures, with the addition of age group and sex. Subsequent statistical commands were run on the imputed data, with the coefficients and standard errors adjusted for the variability between imputations using Rubin's combination rules [26]. Multiple imputation assumes that the missing data is missing-at-random. However, since poor health may be a contributing factor for noncompletion and withdrawal, we included a sensitivity analysis similar to that suggested by Biering and colleagues [27] to see what impact reducing imputed values by 25% had on the model results.

Per-protocol analyses were conducted by repeating these analyses for those respondents who accessed at least 3 cores of the CancerCope intervention and comparing these respondents to the control respondents. Differences in baseline demographic characteristics and baseline measures between respondents who accessed at least 3 cores and those who accessed fewer than 3 cores were analyzed by performing and reporting the results of multivariate backwards stepwise logistic regression analysis.

Effect sizes for the per-protocol analysis were estimated for each continuous outcome variable based on Cohen *d* [28], with the mean difference scores (baseline to 2-month) being compared between the intervention (at least 3 cores accessed) and the patient education group. Test statistics of Cohen *d* and 95% confidence intervals were run for each imputation separately and then combined across the multiple imputations using Rubin's rule [29].

A post-hoc power calculation based on 79 people in the CancerCope intervention and 84 in the patient education arm (163 in total) showed our study cohort provided 89% power to detect a medium effect size (0.5) with a significance level (alpha) of .05 using a 2-sided 2-sample *t* test. All analysis was conducted in Stata 15 (StataCorp LLC).

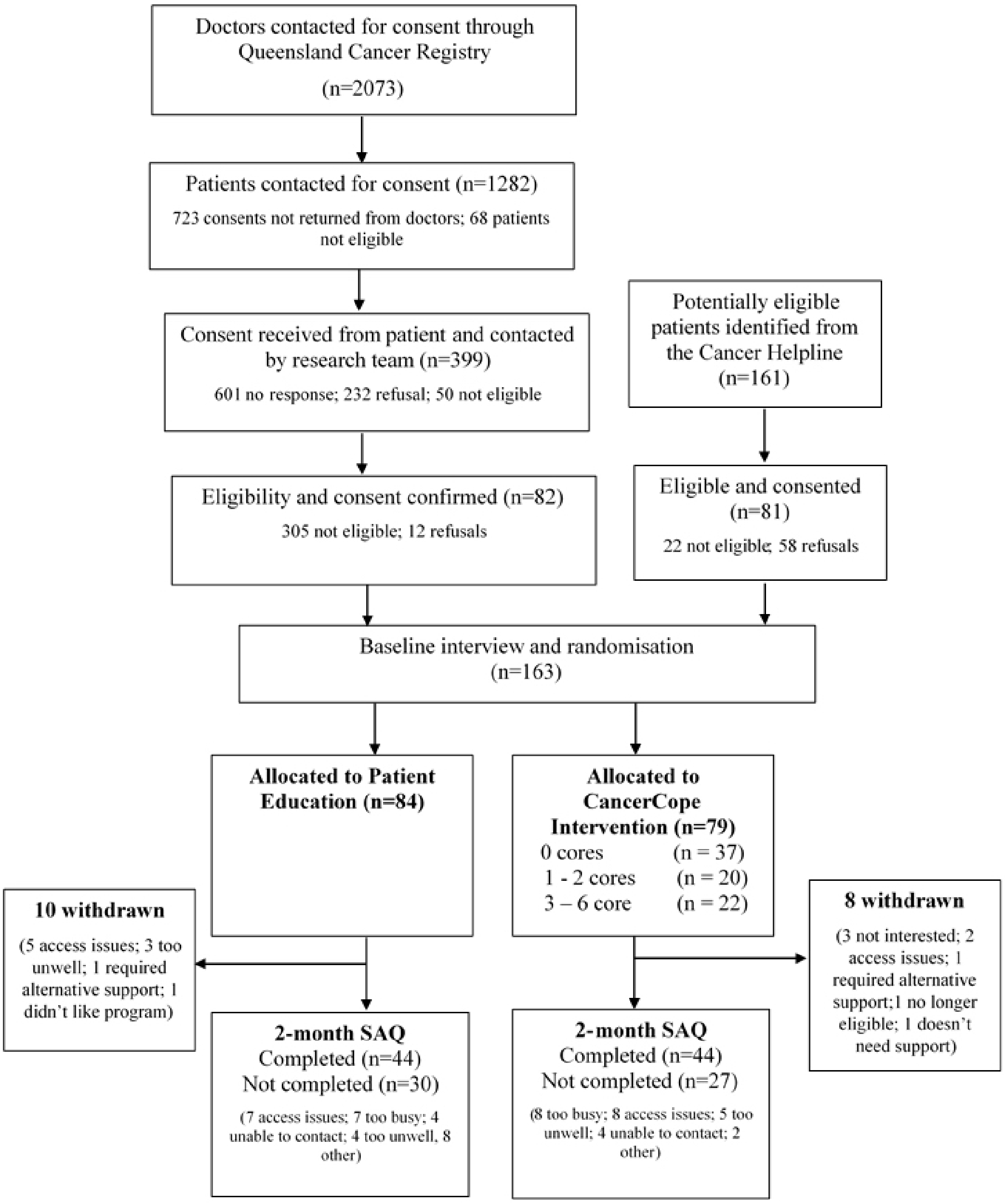
Results

Participants

Between April 2015 and May 2016, a total of 163 participants were recruited through the QCR and the Cancer Helpline (Figure 1) and randomly assigned to patient education (n=84) or CancerCope intervention (n=79). A detailed description of the sample of this trial has been published elsewhere [15]. In the sample, 68.1% (111/163) were female, the mean age of participants was 57 years, over 60% of the sample (100/163) had completed further education after high school, the most common cancer type was colorectal (60/163, 36.8%) followed by breast (42/163, 25.8%) and melanoma (29/163, 17.8%), and median days since diagnosis was 139. Respondents were more likely to complete the second questionnaire if they were retired rather than employed or other work status (χ^2_2=6.8, *P*=.03) or had higher unmet sexuality needs (χ^2_1=5.6, *P*=.02) or lower unmet physical needs (χ^2_1=4.3, *P*=.04).

Of those in the CancerCope intervention group, 10% (8/79) accessed all 6 cores, with 47% (37/79) not accessing any cores; 28% (22/79) accessed 3 or more cores and were classified as completers. For participants in the patient education group, 55% (46/84) accessed the patient education website content. Of those, 61% (28/46) accessed the site once, 33% (15/46) accessed the site 2 to 4 times, and 6% (3/46) accessed the site 5 or more times.

Figure 1. CONSORT flowchart from baseline to 2 months.



Effectiveness

The intention-to-treat analysis (Multimedia Appendix 2) showed no evidence of a statistically significant intervention effect on any of the primary or secondary outcome variables, with these results robust to the missing-at-random assumption (Multimedia Appendix 3). A secondary per-protocol analysis restricted within the CancerCope group to those who accessed at least 3 cores during the study period (n=22) found evidence of a greater decrease in psychological distress (*P*=.03) and cancer-specific distress (*P*=.02) along with unmet psychological needs (*P*=.03) from baseline to 8 weeks compared with the patient education group (Multimedia Appendix 4). Again, these per-protocol results were robust to the missing-at-random assumption (Multimedia Appendix 5).

When comparing the characteristics of patients in the intervention who accessed ≥3 cores against those who accessed fewer than 3 cores, the demographic variables that were retained in the logistic model through the backward selection process were age group (χ^2_2=5.4, *P*=.07), sex (χ^2_1=2.8, *P*=.10), and work status (χ^2_2=9.9, *P*=.01), suggesting that females and younger patients, including younger patients among those who were retired, were slightly more likely to be in the per-protocol group. In addition, there was also some evidence that respondents who had higher unmet information (χ^2_1=2.2, *P*=.14) and patient care (χ^2_1=3.2, *P*=.08) needs, higher cancer-specific distress (χ^2_1=4.7, *P*=.03), and lower posttraumatic growth (χ^2_1=4.0, *P*=.05) were more likely to be in the per-protocol group.

On average, patients in the intervention arm found the CancerCope Web program easy to use, helpful, and a good fit for their needs (Figure 2). The relaxation, meditation, and self-help components were reported as most helpful. Technical problems were infrequent (Figure 3). Patients reported the program as more helpful for improving knowledge, problem solving, and future coping than for mood and would recommend it to others (Figure 4).

Figure 2. Findings from the Internet Evaluation and Utility Questionnaire (from a response of n=41-42).

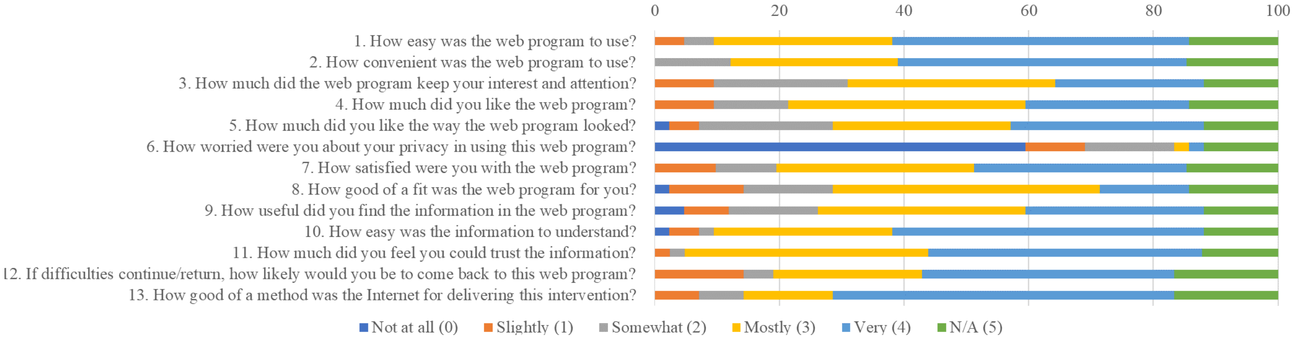


Figure 3. Findings from the Internet Intervention Adherence Questionnaire (from a response of n=40-42).

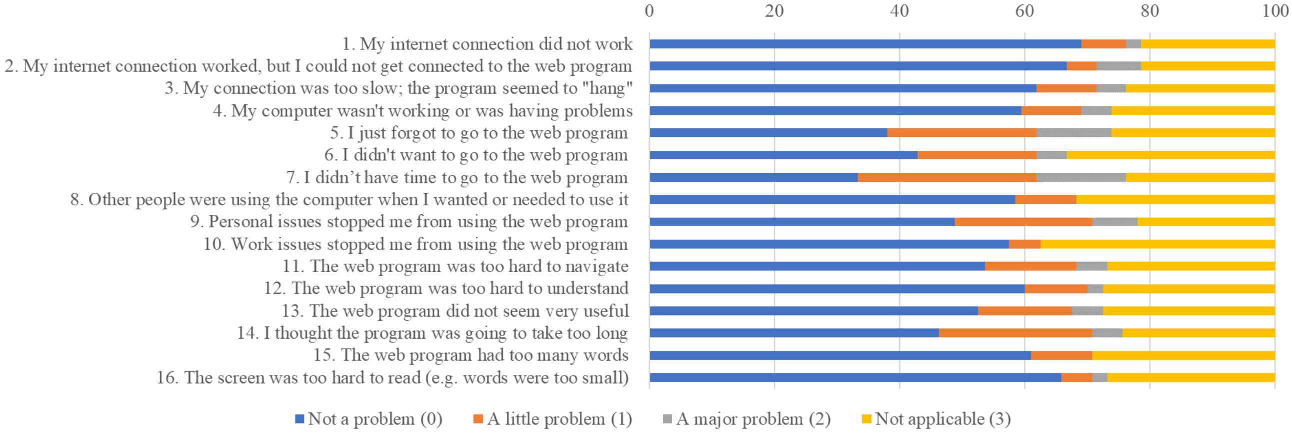
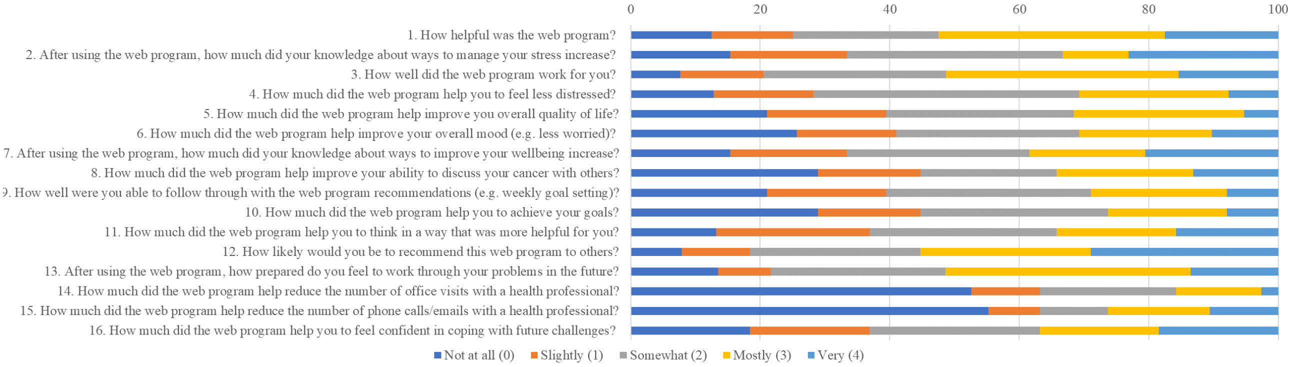


Figure 4. Findings from the Internet Impact and Effectiveness Questionnaire (from a response of n=37-40).



Discussion

Although an intervention effect was not found in the primary analyses, a secondary per-protocol analysis found psychological benefits with medium effects for the subgroup of patients who more closely adhered to the CancerCope program. Hence, while the program overall was positively received by patients, we are not able to conclude it was effective as a standalone psychosocial care intervention. We do, however, have evidence to suggest that if the intervention is used, positive effects can be expected.

The delivery of psychosocial care to cancer patients through a scalable, population-based approach remains an important goal as cancer prevalence increases. The CancerCope program reported in this trial differs from much of the previously reported Web-based psycho-oncology intervention research in that it was a fully automated and tailored intervention and did not include therapist or nurse support or guidance [30], support group forums [31], discussion boards [32,33], or messaging services [34]. Rather, our approach was designed to be completely self-managed by the patient and therefore suitable for widespread dissemination at minimal cost.

One possible way forward may be to view Web-based interventions of this type as an important step in universal psychosocial care within a stepped or tiered model of care [9]. For example, distressed patients or those with unmet psychological care needs could be offered a low-cost self-managed online program such as CancerCope and then stepped or triaged to other more in-depth care models (such as nurse counseling or psychology services) if their distress remains unresolved. Relatedly, and perhaps more efficiently, if we could better identify who might be best served by a Web-based approach as well as who might use the intervention, we could make this type of intervention available to these individuals first. We note that we were not able to recruit our original target sample size and this precluded us from being able to more deeply elucidate the patient subgroups for whom CancerCope was helpful. This is a study limitation. We have previously shown that background variables such as educational level and age moderate the effectiveness of tele-based psychological intervention [35]. Sociodemographic variables such as these may well have influenced participants' responses to this Web-based intervention; however, our study was not able to clearly examine this possibility. Moving forward, we suggest psychosocial researchers and practitioners in cancer care might consider Web-based programs as a component of stepped care and focus further on what works best and for whom.

Acknowledgments

We gratefully acknowledge the support of Dr Cindy Mervin, Dr Lizbeth Kenny, Dr Ian Irving, and Dr Euan Walpole as associate investigators and Ms Leonie Young as a consumer advisor. This project was funded by the Australian National Health and Medical Research Council and Cancer Council Queensland (ID APP1056756). PY is funded by a National Health and Medical Research Council Early Career Fellowship (ID APP1054038). The funding body had no involvement in the design of the study; collection, analysis, and interpretation of data; or writing the manuscript.

Authors' Contributions

SKC, JD, LR, and FT developed the study concept and aims and initiated the project. All other authors assisted in further development of the protocol. SKC was responsible for drafting the manuscript. SKC, SC, and LN implemented the protocol and oversaw collection of the data. PB undertook the data analyses. All authors contributed to the final manuscript.

Conflicts of Interest

LR and FT have equity ownership in BeHealth Solutions, LLC, the company that developed the CancerCope intervention as well as other similar products related to the research reported in this publication. Specifically, BeHealth Solutions, LLC, has the software platform on which CancerCope was built from the University of Virginia. The terms of this arrangement have been reviewed and approved by the University of Virginia in accordance with its conflict of interest policy.

Multimedia Appendix 1

Screenshots of the intervention.

[PDF File (Adobe PDF File), 2MB - jmir_v20i1e42_app1.pdf]

Multimedia Appendix 2

Intention-to-treat analysis (baseline vs 2-month) for primary and secondary outcome scores using multiple imputation analysis (50 imputations).

[PDF File (Adobe PDF File), 137KB - jmir_v20i1e42_app2.pdf]

Multimedia Appendix 3

Sensitivity analysis (25% reduction in imputed values): intention-to-treat analysis (baseline vs 2-month) for primary and secondary outcome scores using multiple imputation analysis (50 imputations).

[PDF File (Adobe PDF File), 138KB - jmir_v20i1e42_app3.pdf]

Multimedia Appendix 4

Per-protocol analysis: intention-to-treat analysis (baseline vs 2-month) for primary and secondary outcome scores using multiple imputation analysis (50 imputations).

[PDF File (Adobe PDF File), 138KB - jmir_v20i1e42_app4.pdf]

Multimedia Appendix 5

Sensitivity analysis (25% reduction in imputed values): per-protocol analysis (baseline vs 2-month) for primary and secondary outcome scores using multiple imputation analysis (50 imputations).

[PDF File (Adobe PDF File), 139KB - jmir_v20i1e42_app5.pdf]

Multimedia Appendix 6

CONSORT‐EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 9MB - jmir_v20i1e42_app6.pdf]

References

1. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Esser S, Mathers C, et al. Cancer incidence and mortality worldwide: IARC CancerBase—GLOBOCAN 2012 v1. URL: http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx [accessed 2017-10-23] [WebCite Cache ID 6uPurqQUx]
2. Zabora J, BrintzenhofeSzoc K, Curbow B, Hooker C, Piantadosi S. The prevalence of psychological distress by cancer site. Psychooncology 2001 Feb;10(1):19-28. [Medline: 11180574]
3. Chambers SK, Ng SK, Baade P, Aitken JF, Hyde MK, Wittert G, et al. Trajectories of quality of life, life satisfaction, and psychological adjustment after prostate cancer. Psychooncology 2017 Jan 11;26(10):1576-1585. [doi: 10.1002/pon.4342] [Medline: 27943512]
4. Dunn J, Ng SK, Breitbart W, Aitken J, Youl P, Baade PD, et al. Health-related quality of life and life satisfaction in colorectal cancer survivors: trajectories of adjustment. Health Qual Life Outcomes 2013 Mar 14;11:46 [FREE Full text] [doi: 10.1186/1477-7525-11-46] [Medline: 23497387]
5. Dunn J, Ng SK, Holland J, Aitken J, Youl P, Baade PD, et al. Trajectories of psychological distress after colorectal cancer. Psychooncology 2013 Aug;22(8):1759-1765. [doi: 10.1002/pon.3210] [Medline: 23125004]
6. Holland J, Watson M, Dunn J. The IPOS new International Standard of Quality Cancer Care: integrating the psychosocial domain into routine care. Psychooncology 2011 Jul;20(7):677-680. [doi: 10.1002/pon.1978] [Medline: 21710503]
7. Schofield P, Chambers S. Effective, clinically feasible and sustainable: key design features of psycho-educational and supportive care interventions to promote individualised self-management in cancer care. Acta Oncol 2015 May;54(5):805-812. [doi: 10.3109/0284186X.2015.1010016] [Medline: 25813474]
8. Chambers SK, Girgis A, Occhipinti S, Hutchison S, Turner J, McDowell M, et al. A randomized trial comparing two low-intensity psychological interventions for distressed patients with cancer and their caregivers. Oncol Nurs Forum 2014 Jul 01;41(4):E256-E266. [doi: 10.1188/14.ONF.E256-E266] [Medline: 24969260]
9. Hutchison SD, Steginga SK, Dunn J. The tiered model of psychosocial intervention in cancer: a community based approach. Psychooncology 2006 Jun;15(6):541-546. [doi: 10.1002/pon.973] [Medline: 16331595]

10. Leykin Y, Thekdi SM, Shumay DM, Muñoz RF, Riba M, Dunn LB. Internet interventions for improving psychological well-being in psycho-oncology: review and recommendations. Psychooncology 2012 Sep;21(9):1016-1025 [FREE Full text] [doi: 10.1002/pon.1993] [Medline: 21608075]
11. 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.
12. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of Web-based vs. non-Web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40 [FREE Full text] [doi: 10.2196/jmir.6.4.e40] [Medline: 15631964]
13. Holland JC, Andersen B, Breitbart WS, Buchmann LO, Compas B, Deshields TL, et al. NCCN clinical practice guidelines in oncology: distress management, version 2.2013. J Natl Compr Canc Netw 2013 Feb 14;11(2):190-209. [doi: 10.6004/jnccn.2013.0027]
14. Hutchison SD, Sargeant H, Morris BA, Hawkes AL, Clutton S, Chambers SK. A community-based approach to cancer counselling for patients and carers: a preliminary study. Psychooncology 2011 Aug;20(8):897-901. [doi: 10.1002/pon.1786] [Medline: 20878875]
15. Chambers SK, Ritterband L, Thorndike F, Nielsen L, Aitken JF, Clutton S, et al. A study protocol for a randomised controlled trial of an interactive web-based intervention: CancerCope. BMJ Open 2017 Jun 23;7(6):e017279 [FREE Full text] [doi: 10.1136/bmjopen-2017-017279] [Medline: 28645985]
16. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann Intern Med 2001 Apr 17;134(8):663-694. [Medline: 11304107]
17. Derogatis L, Lopez M. Brief Symptom Inventory 18: Administration, Scoring and Procedures Manual. Minneapolis: National Computer Systems Inc; 2000.
18. Horowitz M, Wilner N, Alvarez W. Impact of Event Scale: a measure of subjective stress. Psychosom Med 1979 May;41(3):209-218. [Medline: 472086]
19. Weiss D, Marmar C. The Impact of Event Scale–Revised. In: Wilson JP, Keane TM, editors. Assessing Psychological Trauma and PTSD. New York: Guilford Press; 1997:399-411.
20. Boyes A, Girgis A, Lecathelinais C. Brief assessment of adult cancer patients' perceived needs: development and validation of the 34-item Supportive Care Needs Survey (SCNS-SF34). J Eval Clin Pract 2009 Aug;15(4):602-606. [doi: 10.1111/j.1365-2753.2008.01057.x] [Medline: 19522727]
21. Tedeschi RG, Calhoun LG. The Posttraumatic Growth Inventory: measuring the positive legacy of trauma. J Trauma Stress 1996 Jul;9(3):455-471. [Medline: 8827649]
22. Richardson J, Iezzi A, Khan MA, Maxwell A. Validity and reliability of the Assessment of Quality of Life (AQoL)-8D multi-attribute utility instrument. Patient 2014;7(1):85-96 [FREE Full text] [doi: 10.1007/s40271-013-0036-x] [Medline: 24271592]
23. Ritterband LM, Ardalan K, Thorndike FP, Magee JC, Saylor DK, Cox DJ, et al. Real world use of an Internet intervention for pediatric encopresis. J Med Internet Res 2008 Jun 30;10(2):e16 [FREE Full text] [doi: 10.2196/jmir.1081] [Medline: 18653440]
24. Thorndike FP, Saylor DK, Bailey ET, Gonder-Frederick L, Morin CM, Ritterband LM. Development and perceived utility and impact of an Internet intervention for insomnia. E J Appl Psychol 2008;4(2):32-42 [FREE Full text] [Medline: 20953264]
25. Ritterband LM, Borowitz S, Cox DJ, Kovatchev B, Walker LS, Lucas V, et al. Using the internet to provide information prescriptions. Pediatrics 2005 Nov;116(5):e643-e647. [doi: 10.1542/peds.2005-0404] [Medline: 16263978]
26. Rubin D. Multiple Imputation for Nonresponse in Surveys. Hoboken: John Wiley & Sons; 1987.
27. Biering K, Hjollund NH, Frydenberg M. Using multiple imputation to deal with missing data and attrition in longitudinal studies with repeated measures of patient-reported outcomes. Clin Epidemiol 2015;7:91-106 [FREE Full text] [doi: 10.2147/CLEP.S72247] [Medline: 25653557]
28. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd Edition. Hillsdale: Lawrence Earlbaum Associates; 1988.
29. Rubin DB. Inference and missing data. Biometrika 1976;63(3):581-592. [doi: 10.1093/biomet/63.3.581]
30. Ruland CM, Andersen T, Jeneson A, Moore S, Grimsbø GH, Børøsund E, et al. Effects of an internet support system to assist cancer patients in reducing symptom distress: a randomized controlled trial. Cancer Nurs 2013;36(1):6-17. [doi: 10.1097/NCC.0b013e31824d90d4] [Medline: 22495503]
31. Wootten AC, Abbott JM, Meyer D, Chisholm K, Austin DW, Klein B, et al. Preliminary results of a randomised controlled trial of an online psychological intervention to reduce distress in men treated for localised prostate cancer. Eur Urol 2015 Sep;68(3):471-479. [doi: 10.1016/j.eururo.2014.10.024] [Medline: 25454611]
32. Owen JE, Klapow JC, Roth DL, Shuster JL, Bellis J, Meredith R, et al. Randomized pilot of a self-guided internet coping group for women with early-stage breast cancer. Ann Behav Med 2005 Aug;30(1):54-64. [doi: 10.1207/s15324796abm3001_7] [Medline: 16097906]
33. Duffecy J, Sanford S, Wagner L, Begale M, Nawacki E, Mohr DC. Project onward: an innovative e-health intervention for cancer survivors. Psychooncology 2013 Apr;22(4):947-951 [FREE Full text] [doi: 10.1002/pon.3075] [Medline: 22438297]

34. Yun YH, Lee KS, Kim YW, Park SY, Lee ES, Noh DY, et al. Web-based tailored education program for disease-free cancer survivors with cancer-related fatigue: a randomized controlled trial. J Clin Oncol 2012 Apr 20;30(12):1296-1303 [FREE Full text] [doi: 10.1200/JCO.2011.37.2979] [Medline: 22412149]
35. Chambers SK, Ferguson M, Gardiner RA, Aitken J, Occhipinti S. Intervening to improve psychological outcomes for men with prostate cancer. Psychooncology 2013 May;22(5):1025-1034. [doi: 10.1002/pon.3095] [Medline: 22549800]

Abbreviations

CBT: cognitive behavioral therapy
QCR: Queensland Cancer Registry

Edited by G Eysenbach; submitted 29.08.17; peer-reviewed by D Meyer, IW Pulantara; comments to author 19.10.17; revised version received 08.11.17; accepted 16.11.17; published 31.01.18.

Please cite as:
Chambers SK, Ritterband LM, Thorndike F, Nielsen L, Aitken JF, Clutton S, Scuffham PA, Youl P, Morris B, Baade PD, Dunn J
Web-Delivered Cognitive Behavioral Therapy for Distressed Cancer Patients: Randomized Controlled Trial
J Med Internet Res 2018;20(1):e42
URL: http://www.jmir.org/2018/1/e42/
doi:10.2196/jmir.8850
PMID:29386173

Original Paper

A Patient-Held Smartcard With a Unique Identifier and an mHealth Platform to Improve the Availability of Prenatal Test Results in Rural Nigeria: Demonstration Study

Semiu Olatunde Gbadamosi[1], MBBS, MPH; Chuka Eze[2], MSc; John Olajide Olawepo[3], MBBS, MSc; Juliet Iwelunmor[4], PhD; Daniel F Sarpong[5], PhD; Amaka Grace Ogidi[6], MEd; Dina Patel[1], RN, PNP, MSN; John Okpanachi Oko[3], MBBS; Chima Onoka[6], MBBS, MPH, PhD; Echezona Edozie Ezeanolue[1], MPH, MD, FAAP, FIDSA

[1]Global Health Initiative, School of Community Health Sciences, University of Nevada, Las Vegas, Las Vegas, NV, United States
[2]Vitira Health, Arlington, VA, United States
[3]Caritas Nigeria, Abuja, Nigeria
[4]Department of Behavioral Science and Health Education, St Louis University, St Louis, MO, United States
[5]Center for Minority Health and Health Disparities Research and Education, Xavier University, New Orleans, LA, United States
[6]University of Nigeria, Nsukka, Nigeria

Corresponding Author:
Semiu Olatunde Gbadamosi, MBBS, MPH
Global Health Initiative
School of Community Health Sciences
University of Nevada, Las Vegas
4505 S. Maryland Parkway
Las Vegas, NV, 89154-1026
United States
Phone: 1 702 895 4950
Fax: 1 702 895 5573
Email: semiu.gbadamosi@unlv.edu

Abstract

Background: Community-based strategies to test for HIV, hepatitis B virus (HBV), and sickle cell disease (SCD) have expanded opportunities to increase the proportion of pregnant women who are aware of their diagnosis. In order to use this information to implement evidence-based interventions, these results have to be available to skilled health providers at the point of delivery. Most electronic health platforms are dependent on the availability of reliable Internet connectivity and, thus, have limited use in many rural and resource-limited settings.

Objective: Here we describe our work on the development and deployment of an integrated mHealth platform that is able to capture medical information, including test results, and encrypt it into a patient-held smartcard that can be read at the point of delivery without the need for an Internet connection.

Methods: We engaged a team of implementation scientists, public health experts, and information technology specialists in a requirement-gathering process to inform the design of a prototype for a platform that uses smartcard technology, database deployment, and mobile phone app development. Key design decisions focused on usability, scalability, and security.

Results: We successfully designed an integrated mHealth platform and deployed it in 4 health facilities across Benue State, Nigeria. We developed the Vitira Health platform to store test results of HIV, HBV, and SCD in a database, and securely encrypt the results on a Quick Response code embedded on a smartcard. We used a mobile app to read the contents on the smartcard without the need for Internet connectivity.

Conclusions: Our findings indicate that it is possible to develop a patient-held smartcard and an mHealth platform that contains vital health information that can be read at the point of delivery using a mobile phone-based app without an Internet connection.

Trial Registration: ClinicalTrials.gov NCT03027258; https://clinicaltrials.gov/ct2/show/NCT03027258 (Archived by WebCite at http://www.webcitation.org/6owR2D0kE)

(J Med Internet Res 2018;20(1):e18) doi:10.2196/jmir.8716

KEYWORDS

mHealth; prenatal screening; HIV; hepatitis B; sickle cell disease; Nigeria; telemedicine; prenatal diagnosis; infectious disease transmission, vertical

Introduction

Background

Despite significant investment and efforts to address maternal and child health challenges in Nigeria, the country still has one of the highest rates of child mortality in the world, with 108 deaths among children less than 5 years of age per 1000 live births [1]. An estimated 240,000 of the 750,000 children who die before their fifth birthday are newborns. Most of these infant deaths are attributable to preventable infectious diseases, including HIV and hepatitis B virus (HBV) infections, and complications of sickle cell disease (SCD)—the commonest genetic disease [1,2]. Even though programs to prevent mother-to-child transmission (MTCT) of HIV in Nigeria have been expanded, an estimated 41,000 infants became infected with HIV in 2015 [3]. The risk of perinatal transmission of HIV is increased when an HIV-infected pregnant woman is co-infected with HBV, which remains endemic in Nigeria [4,5]. Further, children with SCD are at increased risk of HIV infection due to frequent blood transfusion [6]. In Nigeria, about 50% to 80% of children with SCD die before their fifth birthday due to several complications from the disease [7,8]. Despite the availability of simple, inexpensive interventions to prevent MTCT of HIV and HBV and to manage SCD, including antiretroviral prophylaxis, HBV vaccine, and penicillin prophylaxis, respectively, implementation remains inconsistent as a result of limited availability of diagnostic information at the point of delivery. It is therefore critical not only that pregnant women are screened for these conditions, but also that there is an efficient and effective way of ensuring these results are available at the point of delivery.

In our previous study, we had demonstrated the feasibility of using a community-based, integrated approach to increase uptake of screening for conditions such as HIV infection, HBV infection, and sickle cell genotype among pregnant women in southeastern Nigeria [9,10]. Our Healthy Beginning Initiative (HBI) trial was a cluster-randomized trial designed to reduce barriers to screening for these diseases during the prenatal period. We conducted the HBI trial among pregnant women and their male partners in communities where they resided. The scale-up of HBI in communities across Nigeria has expanded opportunities to increase the proportion of pregnant women who are tested and aware of their diagnosis. In order to use this information to implement evidence-based interventions for the prevention of MTCT of HIV and HBV and management of the infant with SCD, these test results need to be available to a skilled birth attendant at the point of delivery.

Medical record keeping is an arduous task in resource-limited settings. Efforts to implement electronic medical record (EMR) systems to improve collection and availability of and access to personal health information have been met with challenges. Caritas Nigeria, a US President's Emergency Plan for AIDS Relief (PEPFAR)-supported implementing partner providing HIV services in Nigeria, runs an EMR system anchored in the IQCare package managed by the Palladium Group (formerly Futures Group). IQCare is deployed in locations in Nigeria, Uganda, Zimbabwe, and Kenya [11]. It is a robust and customizable database with an SQL back end and an Internet Explorer-based form, query, and report front end. It was to be configured and customized to suit either a paperless or a paper-based setup. The paperless operation mode of IQCare is the ideal setup where computers or other data collection devices are present at each data collection point and are all linked to a network server where the database resides. Electronic versions of the data collection forms are made available on these remote computers or devices such that data are captured seamlessly as part of the business process and uploaded directly into the central server database. For instance, when a patient enrolls at the medical record unit, the data are automatically available at the clinician's office. A prescription made by the clinician can be accessed by the pharmacy in real time. However, limited infrastructure, such as poor Internet connectivity and unreliable electric power, make the paper-based mode more prevalent in this setting. The current practice is for data to be collected on structured paper-based forms and subsequently entered at one central data entry location on-site. This process creates an inefficient system whereby the information may not be easily retrievable and readily available, particularly when needed in an emergency situation such as at the point of delivery.

Specific mHealth solutions targeted at addressing this gap may offer promising opportunities to improve the provision of health services. Recent advancements in mHealth have the potential to strengthen the provision of health services, particularly in resource-limited settings. These advancements include the widespread use of mobile phones and growth in coverage of mobile cellular networks; the rapid rise in the development of mHealth apps and wearable devices; and the availability of higher data transmission speeds around the globe. Despite these advancements, low- and middle-income countries still face major challenges with implementing cost-effective, culturally acceptable, and sustainable mHealth solutions that can be integrated into health systems [12]. The use of low-cost, widely accessible mobile technologies offers an opportunity to augment the effort to digitize record keeping and ensure that vital test results are available to clinicians at the point of care.

Objective

In this paper, we describe our work on the development and deployment of an integrated mHealth platform that is able to capture vital health information, including test results for HIV, HBV, and SCD, and encrypt it into a patient-held smartcard that can be read at the point of delivery without the need for an Internet connection.

Our rationale for focusing on HIV, HBV, and SCD for mHealth is based on several factors: (1) we found high prevalences of HIV, hepatitis B surface antigen, and sickle cell trait of 2% [10], 5%, and 22% [13], respectively, in pregnant women in the communities where HBI was implemented; (2) an integrated approach to screen for multiple diseases rather than a single condition was widely acceptable by the community; and (3) the integration of both infectious and chronic diseases highlights the double burden facing sub-Saharan Africa.

Methods

Ethical Consideration

This study was approved by the Institutional Review Board of the University of Nevada, Las Vegas, NV, USA, and the Nigerian National Health Research Ethics Committee. This study was registered with ClinicalTrails.gov (ClinicalTrials.gov identifier NCT03027258).

Study Design and Settings

The study design details have been previously described [14]. We conducted the study in Benue State, north-central Nigeria. In 2012, Benue State was estimated to have a total population of 5,138,531, of whom 49.6% were female [15]. Its population is predominantly rural, and most are farmers [16]. According to the Nigeria Demographic and Health Survey, only 60% of pregnant women received prenatal care from a skilled provider for the most recent birth [17] and less than half of all deliveries in the state were attended by a skilled birth attendant [15]. Preliminary data from our ongoing US National Institutes of Health-funded study (grant no. R01HD075050; multiple principal investigators: EEE, CO) demonstrate that in Benue State, prevalence rates of HIV, HBV, and sickle cell trait among pregnant women are 7.8%, 11.1%, and 19.1%, respectively. Pregnant women are not routinely screened for HBV infection and sickle cell genotype during prenatal care in Benue State.

Study Procedure

Predevelopment

In February 2016, we assembled an interdisciplinary group of experts to identify the core needs of the platform given the current challenges identified in the provision of health services in the settings. Expert consultations were conducted by (1) implementation science researchers in the United States with extensive work experience in Nigeria, (2) public health consultants, (3) computer scientists and programmers, and (4) a cadre of health workers with HIV programmatic experience based in Nigeria. The group met biweekly via videoconference during the initial stages to review evidence and share experience in concept and design to inform the development of the platform. Our objective was to develop a decentralized medical health records platform that allowed any authorized health professional to, within seconds, retrieve vital personalized health information needed to improve health outcomes. In addition, we sought to develop a solution that could provide patients with more control over their medical health record. We were also constrained by additional requirements. Our solution needed to be low cost and potentially work on any Android-based mobile phone without the need for an external Internet connection or additional physical hardware. On the basis of our constraints, we arrived at some core design and architectural decisions. Our key design decisions focused on usability, scalability, and security. Every aspect of our solution had to be capable of being deployed to new users without extensive training. Knowledge of how to use a basic mobile phone would be sufficient for a health official to download the mobile app, launch the app, scan a patient card, and use the data to effectively manage a patient. At various points during the development of both the mobile app and the back-end database, we obtained feedback from our research teams and users on the design, usability, and effectiveness of our solution.

Development

The mobile app and back-end database were built using various mobile and Web app software development tools. They include the JavaScript jQuery version 3.2.1 and PHP programming languages, Ionic Mobile Development Framework version 3.6.1 (Ionic), Android Software Development Kit version 26.0.1 (Google Inc), Apple Developer Tools Xcode version 7.3.1 (Apple Inc), and MariaDB version 10.2 (MariaDB Foundation) for secure database management. All mobile and Web-based back-end app code was written from scratch and was not preconfigured using an existing open-source app platform. We estimate that 96% of our mobile app code was portable between Android and iOS devices form factors (mobile phone and tablet). On the back-end database server, we developed and deployed a virtual private cloud in the Amazon Web Services (Amazon Web Services, Inc) cloud environment to securely host the back-end app. We integrated Quick Response (QR) code technology as our primary data transport medium. We quickly focused on developing our first minimum viable product. The first working prototype achieved the ability for an authorized medical health official to use the Vitira Health app to identify a patient's medical health information within 20 seconds without the need for an external data connection and EMR solution. We followed a continuous development approach for the development of the rest of the platform. Understanding the importance of patient privacy, we developed and incorporated an algorithm to encrypt patient data on the smartcard and added user authentication capabilities into the mobile app. We put a tremendous amount of effort into the development of the Vitira Health back-end database where patient data were securely managed. When a patient card was printed, the data were encrypted, encoded, and printed onto a smartcard. We also developed the ability to pull, encrypt, and encode patient data from third-party partner EMR systems through secure application programming interfaces. We built both the mobile app and the Web back-end database to be fully scalable and deployable across multiple mobile phone device and Web architectures.

Predeployment

We developed a manual to help facilitate training sessions for end users and outlined defined tasks and performance objectives. A reference manual with module content delivered in screenshots and picture aids was also developed to guide users in navigating through the different menus and options of the platform. We identified and selected health facilities to deploy the Vitira Health mHealth platform based on several criteria. The facility had to (1) receive funding support from PEPFAR through Caritas Nigeria, (2) be a comprehensive treatment facility, and (3) have records of a high volume of HIV-infected patients and deliveries in the preceding year. Comprehensive treatment facility sites offer free HIV testing services, antiretroviral therapy for both adults and children, and services for prevention of MTCT.

Deployment

Hospital administrators at 4 health facilities (Father Matthias Hospital, Naka; Nongu u Kristu ke Sudan hen Tiv Comprehensive Health Center, Garagbohol; Nongu u Kristu ke Sudan hen Tiv Health Center, Uchi; and Mimidoo Clinic, Gungul) gave their consent to participate in the study. Health workers including birth attendants within the health facilities were selected by their respective hospital administrator to participate in a training session. SOG and JOO facilitated a 3-hour training session in each health facility. During the introductory session, participants were provided with a brief overview, key features and capabilities, and a demonstration of a prototype of the mHealth platform. Participants were then given a list of several tasks to perform to expose them to the core features of the prototype. Throughout the session, we obtained and documented valuable informal feedback from the participants and later adjusted our designs accordingly. At the end of the session, a reference manual was given to the health facility.

Results

Development and Deployment of the Integrated mHealth Platform

Between February and August 2016, we successfully developed an integrated mHealth platform—a point-of-care technology solution that incorporates a Web-based database, smartcard technology, and a mobile app. From September to October 2016, we deployed the Vitira Health platform at the 4 health facilities.

The Vitira Health platform was designed specifically to function in remote locations with low Internet availability. We put security at the forefront of our design decisions. Health data had to be encrypted in storage and during transmission. We implemented those controls with encryption mechanisms on the smartcards, mobile app, and app database servers.

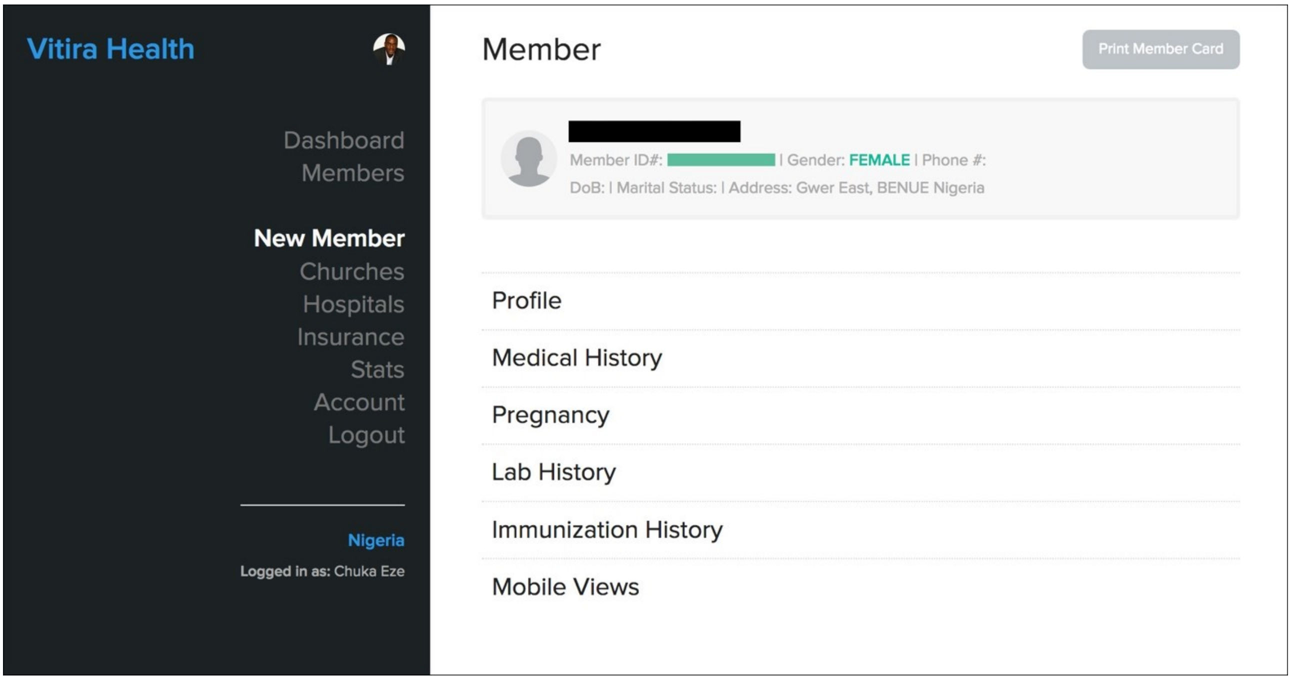
Vitira Health Web-Based Database

On the Vitira Health Web-based back-end platform, we developed the ability for authorized administrators to easily add new patient records, search patient data, and add authorized mobile app users. We also developed capabilities to perform data analytics based on information gathered from deployed apps and health records. Collected patient data, which included demographics and personal health information (Figure 1), were intelligently stored based on unique project categories such as location, disease condition, and additional features that allowed patient records to be pulled into specific slots or buckets for grouping.

Authorized skilled birth attendants (nurses and midwives) at participating health facilities were given access to the Web-based database. They were trained and assigned log-in credentials to access the platform. All access and changes to data were monitored and logged at the back end. On the basis of data collected, we could also see when and where patient records were scanned. The mobile app was able to capture date, time, and global positioning system location of patient interactions, along with the purpose of the health visit. Integration with Google Maps (Google Inc) allowed authorized management to view and track patient visits on the back end.

Multilevel access permissions allowed senior management to limit access to specific data based on permissions granted to each local user on the system. All data were backed up and managed on secure servers, at secure hosting facilities.

Figure 1. The Vitira Health Web-based database for collection of patient health information.



Logged in as: Chuka Eze

[redacted]

Member ID#: [redacted] | Gender: FEMALE | Phone #:

DoB: | Marital Status: | Address: Gwer East, BENUE Nigeria

Figure 2. The patient-held smartcard.

My Health Card

Member #

NTH-9385-495

Vitirahealth.com
Issued 0417 | CE2210

Patient-Held Smartcard

Specific patient data, including HIV test result, date of HIV test, HBV test result, and sickle cell genotype result, were converted to an encrypted QR code and embedded on a card (Figure 2). Each smartcard had a unique identifier and was offered to pregnant women. The data encrypted in the QR code were retrieved using a mobile app to scan the smartcard.

Vitira Health Mobile App

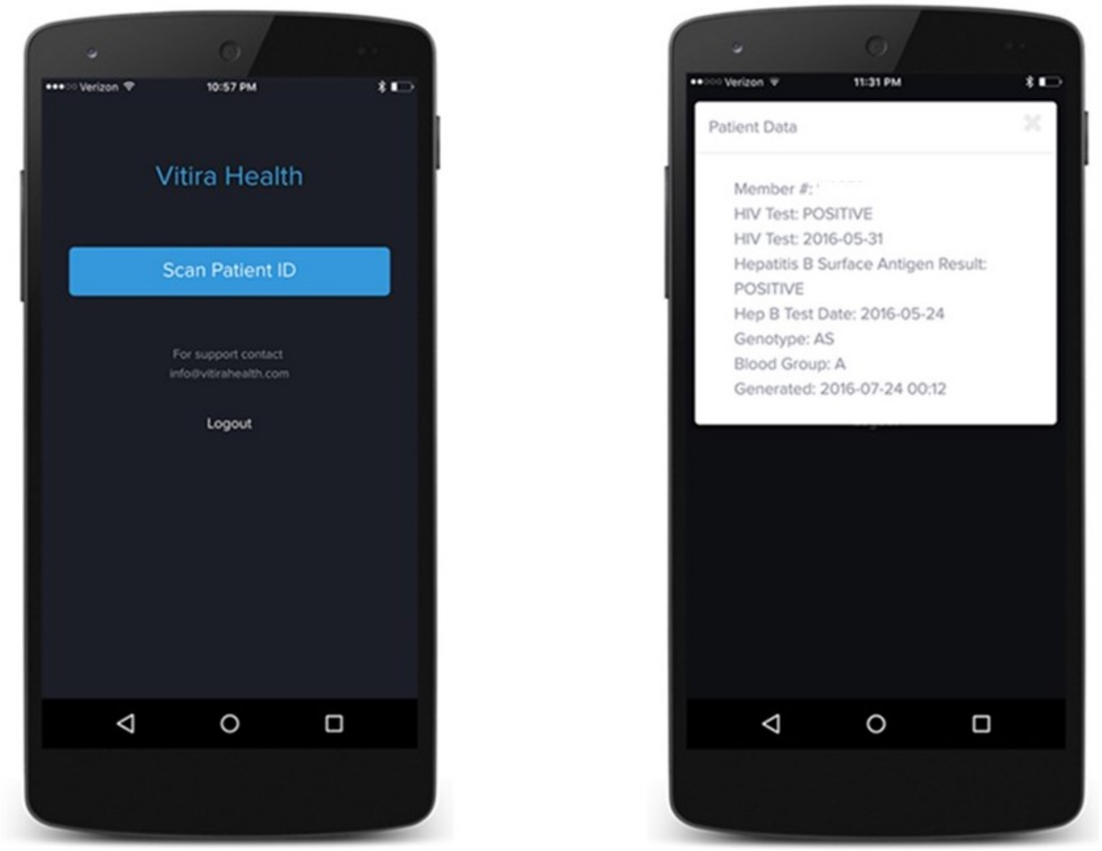
The mobile app was designed as a cross-platform solution to work on Android mobile devices (with a focus on low-end, less-costly Android mobile phones) (Figure 3). We focused heavily on the workflow and sought to have a minimal number of screens and buttons on the app. We developed a modern-looking user interface that was optimized for readability. We implemented the ability for an authorized health official to authenticate once and subsequently gain quick access to the app with a personal identification number code. We developed the Vitira mobile app to work on any camera-enabled mobile phone. Our initial deployment focused on Android mobile phones due to targeted audience factors. While all mobile phones can read our QR code-embedded smartcard, we identified that phones with cameras that were 6 megapixels or higher worked more effectively.

The mobile app was used by authorized skilled birth attendants to scan the smartcard and view patients' data. In doing this, an access log was synchronized to a cloud platform. In addition to scanning a patient's smartcard and viewing the result on the mobile app, the date, time, and location where the smartcard was scanned was synchronized to the database. This was achieved by (1) transmitting the day, time, and location in real time if there was an available Internet connection or (2) storing these data locally on the phone if there was no Internet connectivity and transmitting it later when the connection was available. The entire mHealth platform was built with security in mind such that data across the app were transmitted with 256-bit encryption.

Demographic Characteristics of Health Workers

A total of 19 health workers participated in the training sessions on how to use the mHealth platform. Their mean age was 32.7 (standard deviation [SD] 4.4) years. Most (16/19, 84%) were female and slightly more than half were skilled birth attendants, as Table 1 shows.

Figure 3. The mobile app.



Member #: [redacted]
HIV Test: POSITIVE
HIV Test: 2016-05-31
Hepatitis B Surface Antigen Result: POSITIVE
Hep B Test Date: 2016-05-24
Genotype: AS
Blood Group: A
Generated: 2016-07-24 00:12

Table 1. Demographic characteristics of the health workers using the Vitira Health app (N=19).

Characteristics	Data
Age in years, mean (SD[a])	32.7 (4.4)
Female sex, n (%)	16 (84)
Job description, n	
Birth attendant or community health extension worker	6
Nurse or midwife	5
Clinic director	4
Records or monitoring and evaluation staff	2
Other	2

[a] SD: standard deviation.

Discussion

We were able to demonstrate that we could develop an integrated mHealth platform comprising (1) a secure, Web-based database that contained HIV, HBV, and SCD test results data, (2) a smartcard with an embedded, encrypted, unique QR code, and (3) a mobile app capable of reading the QR code and displaying prenatal test results to authorized skilled birth attendants, even in the absence of Internet connectivity.

Effective interventions provided at the time of childbirth can reduce MTCT of HIV and HBV and guide the management of SCD. These interventions have been shown to reduce the risk of MTCT of HIV to less than 2% [18-20] and have demonstrated 85% to 90% efficacy in preventing MTCT of HBV [21,22]. However, for these interventions to be implemented at the time of childbirth, skilled birth attendants need to know the mother's HIV, HBV, and SCD status. In Nigeria, the health system infrastructure is limited, and EMRs are nonexistent or unreliable due to Internet outages, especially in remote locations [23]. mHealth systems that can provide needed prenatal test results without relying on the Internet are needed.

Nigeria is Africa's largest mobile market with over 150 million mobile phone users and a high penetration of Internet services through mobile networks [24]. The use of mHealth is feasible and potentially can be a cost-effective intervention for improving maternal and perinatal outcomes in Nigeria. It can be used to increase access to health information and reduce turnaround times for receipt of laboratory test results by skilled birth attendants. Studies of mHealth apps to improve maternal health in low-income countries have focused primarily on using mobile phones for data collection, appointment reminders, health promotion and education, and provider-to-provider or person-to-person communication [12,13]. Specifically in Nigeria, mHealth has been used by community health workers for decision support (prenatal care decision support algorithm), health education, and data collection, as well as by clients for appointment reminders [25,26]. To our knowledge, this is the first time that a patient-held smartcard with a unique identifier and QR code specific to the patient has been used. The QR code can be scanned using a mobile phone with an app to access maternal health data, providing needed health information at the point of delivery.

Limitations

In this study, we demonstrated that we could develop an mHealth platform to provide maternal health data using a smartcard and a mobile phone app. Yet we do not know whether patients will (1) accept the smartcard, (2) deliver at one of the health facilities with the mobile phone capable of reading the QR code, or (3) present the smartcard when they arrive at the health facility; nor do we know whether the skilled birth attendants will use this technology to obtain maternal health data. Future research is necessary to determine the acceptability and usability of the mHealth platform on the part of participants and skilled birth attendants. Because this was a demonstration study, we did not measure the impact of mHealth on health outcomes. This is an area for future research.

A limitation to the use of the Vitira Health platform would be the potential for a security breach if unauthorized users gained access to a patient’s smartcard to read the contents of the QR code. However, available QR code readers do not have these capabilities to read the contents due to the embedded security features. We tested commonly used QR code readers publicly available on the Android platform by scanning our smartcard, and all returned scrambled and unreadable data. Also, our mobile app is not available to unauthorized users. A potential breach could occur if an unauthorized user were to gain access to both the log-in credentials and the log-in code to our mobile app to scan a QR code. However, this security breach can easily be tracked because all scans by the mobile app are synchronized to the back end and are monitored by senior-level users. As with most technological solutions, an unreliable power supply may provide challenges that hinder its use. Our mobile app is installed on mobile phone devices that are powered by rechargeable batteries. Compared with the existing system that requires an uninterrupted power supply, mobile phone batteries last longer. Also, the power supply to the mobile phones can be supplemented with power banks that are more affordable than power-generating sets, which also require a constant fuel supply.

Conclusion

We have described our findings on the development of an integrated mHealth platform that can make test results obtained through community-based strategies available at the point of delivery. If future studies demonstrate that the mHealth platform is acceptable and usable by patients and health professionals, this technology could be applied to other chronic health conditions in which previous health data are needed at the point of care to increase the quality and effectiveness of services.

Acknowledgments

Research reported in this paper was supported by the Fogarty International Center of the US National Institutes of Health (NIH) (grant no. R21TW010252) and the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the NIH (grant no. R01HD087994). Additional support for this project was provided by Vitira Health, Caritas Nigeria, the Mapuije Foundation, and the Sunrise Foundation. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. We acknowledge the roles played by Tami Bruno MPH and William Menson MD, MPH during the predevelopment phase of the Vitira Health platform.

Authors' Contributions

EEE, JI, DS, and DP conceived the study. EEE, AGO, JOO, CO, CE, and SOG designed the study. CE provided technical expertise. SOG, CE, and EEE drafted the manuscript. All authors revised and approved the final manuscript.

Conflicts of Interest

CE has equity ownership in Vitira Health, which is developing products related to the research being reported and has applied for a patent.

References

1. Victora CG, Requejo JH, Barros AJD, Berman P, Bhutta Z, Boerma T, et al. Countdown to 2015: a decade of tracking progress for maternal, newborn, and child survival. Lancet 2016 May 14;387(10032):2049-2059. [doi: 10.1016/S0140-6736(15)00519-X] [Medline: 26477328]
2. UNICEF Nigeria. The children: maternal and child health. 2017. URL: https://www.unicef.org/nigeria/children_1926.html [accessed 2017-05-20] [WebCite Cache ID 6wLopWrcy]

3. Joint United Nations Programme on HIV/AIDS (UNAIDS). Prevention gap report. 2016. URL: http://www.unaids.org/sites/default/files/media_asset/2016-prevention-gap-report_en.pdf [accessed 2017-05-20] [WebCite Cache ID 6wLpJllwt]
4. Utoo BT. Hepatitis B surface antigenemia (HBsAg) among pregnant women in southern Nigeria. Afr Health Sci 2013 Dec;13(4):1139-1143 [FREE Full text] [doi: 10.4314/ahs.v13i4.39] [Medline: 24940343]
5. Mbaawuaga E, Iroegbu C, Ike A, Terver G, Jombo A. Studies on prevalence, co-infection and associated risk factors of hepatitis B virus (HBV) and human immunodeficiency virus (HIV) in Benue State, Nigeria. Sci J Public Health 2014;2(6):569-576. [doi: 10.11648/j.sjph.20140206.21]
6. Ugwu RO, Eneh AU. Nonvertical (horizontal) route of HIV transmission in children. J Acquir Immune Defic Syndr 2014 Mar 01;65(3):e128-e130. [doi: 10.1097/01.qai.0000435601.73469.8d] [Medline: 24091695]
7. Fleming AF. The presentation, management and prevention of crisis in sickle cell disease in Africa. Blood Rev 1989 Mar;3(1):18-28. [Medline: 2650773]
8. Piel FB, Hay SI, Gupta S, Weatherall DJ, Williams TN. Global burden of sickle cell anaemia in children under five, 2010-2050: modelling based on demographics, excess mortality, and interventions. PLoS Med 2013 Jul;10(7):e1001484 [FREE Full text] [doi: 10.1371/journal.pmed.1001484] [Medline: 23874164]
9. Ezeanolue EE, Obiefune MC, Yang W, Obaro SK, Ezeanolue CO, Ogedegbe GG. Comparative effectiveness of congregation-versus clinic-based approach to prevention of mother-to-child HIV transmission: study protocol for a cluster randomized controlled trial. Implement Sci 2013 Jun 08;8:62 [FREE Full text] [doi: 10.1186/1748-5908-8-62] [Medline: 23758933]
10. Ezeanolue EE, Obiefune MC, Ezeanolue CO, Ehiri JE, Osuji A, Ogidi AG, et al. Effect of a congregation-based intervention on uptake of HIV testing and linkage to care in pregnant women in Nigeria (Baby Shower): a cluster randomised trial. Lancet Glob Health 2015 Nov;3(11):e692-e700 [FREE Full text] [doi: 10.1016/S2214-109X(15)00195-3] [Medline: 26475016]
11. IQCare. IQCare version 3.6. 2017. URL: http://fgiqcare.codeplex.com/ [accessed 2017-10-07] [WebCite Cache ID 6wLnHfw9R]
12. World Health Organization. mHealth: New Horizons for Health Through Mobile Technologies (Global Observatory for Ehealth). Geneva, Switzerland: World Health Organization; 2011.
13. Burnham-Marusich AR, Ezeanolue CO, Obiefune MC, Yang W, Osuji A, Ogidi AG, et al. Prevalence of sickle cell trait and reliability of self-reported status among expectant parents in Nigeria: implications for targeted newborn screening. Public Health Genomics 2016 Sep;19(5):298-306 [FREE Full text] [doi: 10.1159/000448914] [Medline: 27614873]
14. Ezeanolue EE, Gbadamosi SO, Olawepo JO, Iwelunmor J, Sarpong D, Eze C, et al. An mHealth framework to improve birth outcomes in Benue State, Nigeria: a study protocol. JMIR Res Protoc 2017 May 26;6(5):e100 [FREE Full text] [doi: 10.2196/resprot.7743] [Medline: 28550003]
15. National Bureau of Statistics of Nigeria. Nigeria Social Statistics Atlas: Benue, Nigeria. 2017. URL: http://nigeriahealthatlas.opendataforafrica.org/search?query=benue&source=HomePage [accessed 2017-07-14]
16. Government of Benue State of Nigeria OBSON. Quick Facts: Benue State of Nigeria. 2017. URL: http://benuestate.gov.ng/quick-facts/ [accessed 2018-01-12] [WebCite Cache ID 6wQ11XXf1]
17. National Population Commission, ICF International. Nigeria Demographic and Health Survey. Abuja, Nigeria: NPC; 2013.
18. European Collaborative Study. Mother-to-child transmission of HIV infection in the era of highly active antiretroviral therapy. Clin Infect Dis 2005 Feb 01;40(3):458-465. [doi: 10.1086/427287] [Medline: 15668871]
19. Mofenson LM, McIntyre JA. Advances and research directions in the prevention of mother-to-child HIV-1 transmission. Lancet 2000 Jun 24;355(9222):2237-2244. [doi: 10.1016/S0140-6736(00)02415-6] [Medline: 10881908]
20. Mofenson LM. Successes and challenges in the perinatal HIV-1 epidemic in the United States as illustrated by the HIV-1 Serosurvey of childbearing women. Arch Pediatr Adolesc Med 2004 May;158(5):422-425. [doi: 10.1001/archpedi.158.5.422] [Medline: 15123471]
21. Beasley RP, Hwang LY, Stevens CE, Lin CC, Hsieh FJ, Wang KY, et al. Efficacy of hepatitis B immune globulin for prevention of perinatal transmission of the hepatitis B virus carrier state: final report of a randomized double-blind, placebo-controlled trial. Hepatology 1983;3(2):135-141. [Medline: 6339349]
22. Stevens CE, Toy PT, Tong MJ, Taylor PE, Vyas GN, Nair PV, et al. Perinatal hepatitis B virus transmission in the United States. Prevention by passive-active immunization. JAMA 1985;253(12):1740-1745. [Medline: 3974052]
23. Jimoh L, Pate MA, Lin L, Schulman KA. A model for the adoption of ICT by health workers in Africa. Int J Med Inform 2012 Nov;81(11):773-781. [doi: 10.1016/j.ijmedinf.2012.08.005] [Medline: 22986218]
24. Nigerian Communications Commission. Stakeholder Information. Statistics and Reports. Industry statistics. 2017. URL: https://www.ncc.gov.ng/stakeholder/statistics-reports/industry-overview [WebCite Cache ID 6wQ1F0Vuf]
25. McNabb M, Chukwu E, Ojo O, Shekhar N, Gill CJ, Salami H, et al. Assessment of the quality of antenatal care services provided by health workers using a mobile phone decision support application in northern Nigeria: a pre/post-intervention study. PLoS One 2015;10(5):e0123940 [FREE Full text] [doi: 10.1371/journal.pone.0123940] [Medline: 25942018]
26. Adanikin AI, Awoleke JO, Adeyiolu A. Role of reminder by text message in enhancing postnatal clinic attendance. Int J Gynaecol Obstet 2014 Aug;126(2):179-180. [doi: 10.1016/j.ijgo.2014.02.009] [Medline: 24794692]

Abbreviations

EMR: electronic medical record
HBI: Healthy Beginning Initiative
HBV: hepatitis B virus
MTCT: mother-to-child transmission
PEPFAR: President’s Emergency Plan for AIDS Relief
QR: Quick Response
SCD: sickle cell disease
SD: standard deviation

Edited by G Eysenbach; submitted 14.08.17; peer-reviewed by SE Brown, A Ranjine-Mullings, M Reid, S Abdulrahman, K Claborn; comments to author 15.09.17; revised version received 08.12.17; accepted 02.01.18; published 15.01.18.

Please cite as:
Gbadamosi SO, Eze C, Olawepo JO, Iwelunmor J, Sarpong DF, Ogidi AG, Patel D, Oko JO, Onoka C, Ezeanolue EE
A Patient-Held Smartcard With a Unique Identifier and an mHealth Platform to Improve the Availability of Prenatal Test Results in Rural Nigeria: Demonstration Study
J Med Internet Res 2018;20(1):e18
URL: http://www.jmir.org/2018/1/e18/
doi:10.2196/jmir.8716
PMID:29335234

Original Paper

Leveraging Social Networking Sites for an Autoimmune Hepatitis Genetic Repository: Pilot Study to Evaluate Feasibility

Megan Comerford[1*], BS, MS; Rachel Fogel[1*], BA; James Robert Bailey[2], MD; Prianka Chilukuri[2], MD; Naga Chalasani[1], MD; Craig Steven Lammert[1], MD

[1]Division of Digestive and Liver Diseases, Indiana University School of Medicine, Indianapolis, IN, United States

[2]Department of Medicine, Indiana University School of Medicine, Indianapolis, IN, United States

[*]these authors contributed equally

Corresponding Author:
Craig Steven Lammert, MD
Division of Digestive and Liver Diseases
Indiana University School of Medicine
702 Rotary Circle
Indianapolis, IN,
United States
Phone: 1 317 274 6492
Fax: 1 317 278 6870
Email: clammert@iu.edu

Abstract

Background: Conventional approaches to participant recruitment are often inadequate in rare disease investigation. Social networking sites such as Facebook may provide a vehicle to circumvent common research limitations and pitfalls. We report our preliminary experience with Facebook-based methodology for participant recruitment and participation into an ongoing study of autoimmune hepatitis (AIH).

Objective: The goal of our research was to conduct a pilot study to assess whether a Facebook-based methodology is capable of recruiting geographically widespread participants into AIH patient-oriented research and obtaining quality phenotypic data.

Methods: We established a Facebook community, the Autoimmune Hepatitis Research Network (AHRN), in 2014 to provide a secure and reputable distillation of current literature and AIH research opportunities. Quarterly advertisements for our ongoing observational AIH study were posted on the AHRN over 2 years. Interested and self-reported AIH participants were subsequently enrolled after review of study materials and completion of an informed consent by our study coordinator. Participants returned completed study materials, including epidemiologic questionnaires and genetic material, to our facility via mail. Outside medical records were obtained and reviewed by a study physician.

Results: We successfully obtained all study materials from 29 participants with self-reported AIH within 2 years from 20 different states. Liver biopsy results were available for 90% (26/29) of participants, of which 81% (21/29) had findings consistent with AIH, 15% (4/29) were suggestive of AIH with features of primary biliary cholangitis (PBC), and 4% (1/29) had PBC alone. A total of 83% (24/29) had at least 2 of 3 proposed criteria: positive autoimmune markers, consistent histologic findings of AIH on liver biopsy, and reported treatment with immunosuppressant medications. Self-reported and physician records were discrepant for immunosuppressant medications or for AIH/PBC diagnoses in 4 patients.

Conclusions: Facebook can be an effective ancillary tool for facilitating patient-oriented research in rare diseases. A social media-based approach transcends established limitations in rare disease research and can further develop research communities.

(J Med Internet Res 2018;20(1):e14) doi:10.2196/jmir.7683

KEYWORDS

autoimmune hepatitis; social media; rare disease

Introduction

Challenges in recruiting patients into rare disease clinical studies contribute to wide knowledge gaps in understanding disease pathogenesis, natural history, and optimal therapeutic approaches. In the United States, rare diseases are defined as affecting fewer than 200,000 people at a given time. Despite these seemingly small numbers, there are close to 7000 different rare diseases that impact more than 25 million Americans (7% of the population) [1]. Autoimmune hepatitis (AIH), a chronic liver disease with a prevalence of 24 in 100,000 people [2], is characterized by immune-mediated destruction of hepatic parenchyma, which can result in cirrhosis and death [3-6]. High-impact studies in this field have historically required coordinated multicenter approaches to achieve the participant numbers that provide meaningful results. Obstacles encountered with such studies include increased cost and effort to locate collaborators, successfully train study personnel, and obtain additional ethics approvals [7,8]. Time and money spent on travel by participants can also be barriers to recruitment even for single-site studies. Even when these coordinated approaches are successful, the low density of academic research centers and sparse distribution of patients in North America leaves a large percentage of patients without access to research opportunities to take a more active role in scientific advancement of their own disease [7,9].

The digital age has created new tools for rare disease research, as many proactive patients connect with each other and investigators through online support groups on social networking sites (SNS) such as Facebook and Twitter [10]. With nearly 2 billion daily active users, Facebook is the largest SNS worldwide and consumes 40 minutes of the average American user’s day [11,12]. Individuals in disease-specific groups are motivated to stay up to date with the latest research relevant to their conditions and have even engaged investigators to initiate studies within their respective online groups [13]. Collaboration between such active SNS patient groups and technologically skilled investigators can potentially overcome critical challenges in rare disease research and reduce the need for multicenter approaches.

Considering the established barriers to AIH research and the abundance of AIH patient support groups on SNS, we formed the physician-led Autoimmune Hepatitis Research Network (AHRN) Facebook group to foster patient support, share current AIH information, and promote AIH research efforts. In 2014, we initiated an observational study of patients with AIH (Genetic Repository of Autoimmune Liver Diseases and Contributing Exposures, GRACE Study) for investigating genetic and environmental underpinnings in disease development and clinical outcomes. We conducted a pilot study to assess whether a Facebook-based methodology is capable of recruiting geographically widespread participants into AIH patient-oriented research and obtaining quality phenotypic data.

Methods

Autoimmune Hepatitis Research Network

We established the AHRN on Facebook in 2014 to provide an online repository of summarized AIH literature, discussion of study results and application to patient care, and opportunities to participate in our ongoing AIH research studies. As of May 2017, the AHRN Facebook group has 1640 members and is led and moderated by a hepatologist (CL) from Indiana University (IU). New membership has grown by individuals finding the page via Facebook search results, referrals shared in other AIH Facebook groups by our members, and direct physician-to-clinical-patient marketing. The AHRN is a private Facebook group; individual users wishing to join must be granted approval from the group moderator. Once approved, members are able to view and share content with other group members. A private group setting has the benefit of protecting the privacy of members by hiding their group affiliation from nonmember friends and allows group moderators to screen potential members for illegitimate accounts.

Study Recruitment

We posted a GRACE study recruitment advertisement, consisting of a recruitment message, study summary sheet, and representative study image, on the AHRN Facebook group after obtaining institutional review board approval. The study summary sheet outlined study aims, procedures, inclusion criteria, and the study coordinator’s contact information (phone and email). Inclusion criteria consisted of individuals 18 years and older, diagnosis of AIH or AIH with features of overlap disease (primary biliary cholangitis [PBC] or primary sclerosing cholangitis [PSC]) established by a medical doctor, and no provision of medical care at IU or surrounding affiliate institutions. The study advertisement was posted on the AHRN Facebook group at quarterly intervals over 24 months by the group moderator. The study advertisement directed interested individuals to contact the study coordinator to discuss study objectives and subsequently give consent by phone.

After verbal consent, a study packet containing epidemiologic surveys [14,15], an Oragene DNA saliva kit (DNA Genotek Inc), study instructions, consent documents, health information release form, and prepaid return envelope were mailed to the participant’s home address. Self-reported demographic parameters and medical history were collected from a single survey [14]. Participants were then assigned a study identification number linking all completed materials. Upon receipt of study materials, the participant’s health information release was faxed to the physician treating the AIH in order to collect supportive data related to diagnostic testing. Requested medical records included procedure notes and reports of liver biopsy and pathology and antinuclear antibody, antismooth muscle antibody, antiliver kidney microsomal antibody, antisoluble liver antigen, antimitochondrial antibody (AMA), both cytoplasmic and perinuclear antineutrophil cytoplasmic, immunoglobulin G, and viral hepatitis serologies. Failure to receive a participant’s study materials or outside medical records within 1 month of distribution resulted in repeat contact to each entity up to 3 times. Outside medical record nonresponse was further targeted with a follow-up phone call to the appropriate clerk on the day of medical release transmission.

Data Collection and Analysis

Once received, outside medical records were reviewed and tabulated into the local AIH database by study personnel. Missing clinical and laboratory support data for AIH diagnosis from outside records were identified and subsequently targeted with repeat release of information requests. Collected results (including saliva kit, epidemiologic surveys, and outside diagnostic reports) were stored in a secure location.

DNA Collection and Quantification

Saliva was collected from the subjects using an Oragene collection kit and stored according to kit instructions. DNA was isolated from 500 µL of saliva in PrepIT-L2P Purifier Reagent (Lot PT150219-A, DNA Genotek Inc), and remaining collected saliva sample was saved for future study.

Results

Recruitment Outcomes

We enrolled 29 participants in 2 years between June 2014 and June 2016. Participants were from 20 different states, with 27 (93%) indicating a primary residence outside of Indiana and a median distance of 727 miles from our center. Of 29 participants, 28 (97%) successfully completed and returned study instruments within 3 months of the coordinator-led phone consent. All participants were female, 90% (26/29) were white, and median age at enrollment was 52 years.

Participant Characteristics

Of 28 participants, 20 (71%) subjectively reported a diagnosis of AIH alone, while 29% (8/28) reported AIH with features of PBC (Table 1). Among all included participants, 46% (13/28) reported concurrent, extrahepatic autoimmune disease (data not shown). Other frequent medical conditions and symptoms reported included history of contraceptive use (26/28, 93%), history of multiple urinary tract infections (19/28, 68%), frequent fatigue (19/28, 68%), occasional to constant itching (13/28, 46%), and right upper abdominal discomfort (17/28, 61%).

All participants recorded their current medication list within an epidemiologic survey provided at enrollment (Table 2). Discrepancies were found between the participant surveys and medical records for immunosuppressant medication usage. Additionally, 3 participants reported no active immunosuppressant use, which was contradicted by their outside medical records. One participant had no documented evidence of immunosuppression in the physician notes and also had an incomplete simplified score, high AMA titer, and a biopsy consistent with PBC alone.

Appropriate outside medical records were received after the first request along with the medical information release form for 79% (23/29) of participants (Table 1). A total of 50% (13/26) of available liver biopsies were typical for AIH, 1 (4%) indicated PBC only, and 12 (46%) were identified as compatible with AIH according to the simplified criteria histologic findings [16]. Of the participants with definite AIH per the simplified criteria, 25% (1/4) had a clinically reported a positive AMA titer as well as findings on biopsy of AIH with features of PBC.

Three-quarters of the participants (22/29, 76%) had available autoimmune markers for the simplified score assessment. However, there were a number of missing values for the simplified criteria given the nature of data collection. Therefore, in order to enroll AIH patients into the GRACE study, we sought to categorize the participants most likely to have AIH based on meeting at least 2 of 3 lab- or clinician-reported criteria: positive autoimmune markers (simplified criteria), liver biopsy with typical findings of AIH, and/or reported treatment with immunosuppressant medications. Of 29 participants, 24 (83%) met at least 2 of 3 criteria, with 19 of these participants (79%) satisfying all 3.

DNA Isolation Results

Saliva samples were obtained from all 29 study participants. All collected samples were at least 2 mL in volume. DNA isolation from representative saliva samples yielded sufficient quantity and quality of DNA for genetic investigations. Median value of DNA yield was 41 µg per 500 µL saliva sample (interquartile range [IQR] 41.86) and 260/280 ratio was 1.84 (IQR 0.125).

Table 1. Objective participant data collected from outside medical records.

Study no.	No. of medical record attempts	Simplified AIH[a] score parameters					Immunosuppression per MD[c] note	Diagnosis per biopsy
		Autoantibody score	IgG[b] score	Viral score	Histologic score	Complete simplified score		
1	1	0	2	2	2	6	Yes	AIH
2	3	2	—	2	0	—	No	PBC[d]
3	3	2	—	—	2	—	Yes	AIH
4	3	—	—	2	2	—	Yes	AIH
5	1	2	2	2	2	8	Yes	AIH
6	3	2	—	—	2	—	Yes	AIH
7[e]	3	—	—	—	—	—	Yes	—
8	1	2	—	2	1	—	Yes	AIH
9	2	2	2	2	2	8	Yes	AIH
10	1	2	0	2	2	6	Yes	AIH/PBC
11	3	—	—	—	2	—	Yes	AIH
12	3	—	—	—	1	—	Yes	AIH
13	3	0	—	2	1	—	Yes	AIH
14	1	2	2	2	2	8	Yes	AIH/PBC
15	2	2	0	2	2	6	Yes	AIH
16	1	—	—	0	1	—	—	AIH
17	1	0	0	2	1	3	Yes	AIH
18	1	1	2	2	1	6	Yes	AIH/PBC
19	1	2	0	2	2	6	Yes	AIH
20	1	—	—	2	2	—	Yes	AIH/PBC
21	1	2	0	2	1	5	Yes	AIH
22	1	2	1	2	2	7	Yes	AIH/PBC
23	1	2	—	—	2	—	Yes	AIH
24[e]	1	—	—	—	—	—	—	AIH
25	1	0	0	0	1	1	Yes	AIH
26	1	2	2	2	2	8	Yes	AIH
27	1	2	0	2	2	6	Yes	AIH or AIH/PBC
28	1	1	—	2	1	—	Yes	AIH
29	1	2	—	2	1	—	Yes	AIH

[a] AIH: autoimmune hepatitis.

[b] IgG: immune globulin.

[c] MD: medical doctor.

[d] PBC: primary biliary cholangitis.

[e] Participants without biopsy results available for review.

Table 2. Subjective participant data collected from epidemiologic survey.

Study no.	AIH[a]	PBC[b]	No. of immuno-suppression prescriptions	UDCA[c]
1	Yes	Yes	2	No
2	Yes	Yes	0	Yes
3	Yes	No	0	Yes
4	Yes	No	2	Yes
5	Yes	No	2	Yes
6	Yes	No	2	No
7	Yes	Yes	2	No
8	Yes	No	1	No
9	Yes	No	1	No
10	Yes	Yes	1	Yes
11	Yes	No	2	No
12	Yes	No	0	No
13	Yes	No	0	No
14	Yes	Yes	0	Yes
15	Yes	No	1	Yes
16	Yes	No	0	No
17	Yes	No	1	No
18	Yes	Yes	1	Yes
19	Yes	No	1	Yes
20	Yes	Yes	—	Yes
21	Yes	No	0	No
22	Yes	Yes	1	Yes
23	Yes	No	2	No
24	Yes	No	2	No
25	Yes	No	2	No
26	Yes	No	2	No
27	Yes	Yes	2	No
28	Yes	No	1	No
29	Yes	No	1	No

[a] AIH: autoimmune hepatitis.

[b] PBC: primary biliary cholangitis.

[c] UDCA: ursodeoxycholic acid.

Discussion

Principal Results

Sustained AIH research advances are continually impeded by low participant enrollment numbers, limited and disconnected investigators, and poor engagement of geographically spread patients wanting to contribute to scientific progress [9]. However, social media presents a unique platform to address these significant investigational barriers. Our research engagement with the AHRN Facebook group supports patient-oriented research and facilitates participation in research studies even from a distance. Furthermore, we assert that our research team correctly phenotyped and recruited 24 AIH patients and 4 AIH patients with overlap features of PBC to the GRACE study from outside our institutional reach.

The retrieval of supporting outside records and eventual clinical diagnosis of AIH can be challenging for participants recruited both locally and remotely. Remote access was impeded most frequently by delayed response from corresponding medical record departments. Despite the obstacles inherently imposed by the study design, we were able to successfully collect epidemiologic questionnaires and biologic samples from all 29 participants using Facebook. While the effort spent on retrieval of outside records was time intensive, the passive approach to recruitment required less time spent identifying and recruiting patients enrolled locally through traditional methods. The successful recruitment of 29 subjects was substantial for the GRACE study, as our traditional approach at a large academic center provided only 120 participants (24% more cases obtained with social media) in the same duration of study.

SNS have evolved rapidly over the past decade to support our need for social connectedness and instant digital knowledge. Application of these tools in rare disease research is important, as they transcend well-established limitations such as cost, prolonged study courses, and geographic barriers. Our study was cost- and time-efficient and effective at recruiting patients up to 2200 miles away from our medical center. Furthermore, the cohort was similar to prior work according to gender (female predominant), ethnicity, concurrent autoimmune illnesses, fatigue, and itch [2,17]. Participants in similar studies at our institution are often reimbursed US $25 for expenses related to travel, parking, or time burden. Those expenses were not incurred by the GRACE study participants and thus reduced overall cost of study conduct by approximately US $725. Furthermore, our approach to research also reduces time burden for study participants. Our Web-based approach allows participants to complete the study procedures at their convenience rather than on the schedule of the study team and eliminates the additional time spent traveling to and from the study center.

Limitations

In our study, prospective participants quickly engaged our study coordinator for recruitment. However, self-attestation of AIH diagnosis and current treatment strategies were not always congruent with their physician-composed medical records. Furthermore, we found certain medical records had insufficient documentation or lacked specific workup components. Our assessment of at least 2 of 3 key criteria, as outlined above, suggests AIH diagnosis in 24 out of 29 study participants (83%). Other studies have shown similar levels of agreement between self-reported diagnosis data and medical record documentation using traditional research methods, suggesting this challenge is not unique to social media–based research [18,19].

Successful implementation of a social media–based methodology comes with its own unique set of challenges. While SNS can supplement traditional recruitment, the process of sending and receiving medical records and study materials from affiliated institutions and participants proved to be time-intensive for the study team. Nonetheless, the time cost associated with this methodology is considerably less compared to the effort required to initiate recruitment at multiple study sites to achieve a similar geographical reach. Additionally, the demographics of AIH patients support a social media approach as they align well with the SNS user demographics [20]. This group may be better able to provide accurate self-reported data, as it has been found that female gender and higher educational attainment, characteristics of SNS users, are associated with increased uniformity between patient reports and medical records [19]. The demographic similarity between AIH patients and SNS users may not exist for some rare diseases and must be considered when evaluating the applicability of social media methods to specific patient populations.

Finally, a notable challenge of implementing SNS in medical research is simply its acceptability by providers and health systems. The medical community has been hesitant to embrace the potential of this approach to research recruitment, in part due to the risk of Health Insurance Portability and Accountability Act violations and the potential for demographic biases [9,21]. Furthermore, low acceptance may stem from slow adoption of new technologies, particularly in fast-paced clinical environments, and research regarding the use of SNS in medical research is still limited [22]. The rise of mobile health technologies using mobile phones and wearable devices will create new opportunities for conducting remote clinical studies drawing participants from SNS. We believe as these technologies become more prevalent in clinical and research settings, acceptance of new research methodologies such as SNS will also increase. Failure to integrate this model represents a missed opportunity, as many of the over 2 billion daily Facebook users have actively sought medical information or participated in health-related digital support groups [23,24].

Future Applications and Conclusions

We foresee opportunities for social media to help connect like-minded investigators for research collaborations [25]. As we continue to supplement site-based recruitment for GRACE with Facebook-based participant enrollment, we anticipate additional opportunities for social media to be used to complement traditional studies through study advertisements, recruitment opportunities, collaborations with patient groups, or study-specific groups or pages designed for easy communication with study participants. Using SNS for recruiting participants for research regarding rare diseases such as AIH serves as a viable bridge to connect patients to investigators despite geographical barriers and a lack of local opportunities. We are confident our method could become a frontrunner in the improvement of low study numbers in rare disease research. In overcoming recruitment limitations, we now seek a way to efficiently expedite the time between recruiting individuals, obtaining their biologic samples and epidemiological data, and confirming their diagnosis.

In summary, SNS can be an effective tool for facilitating patient-oriented research in AIH. We have successfully implemented a Facebook-based methodology to engage individuals unable to participate in research through traditional methods due to geographic or financial barriers. Mindful use of SNS can transcend established limitations in rare disease research and further cultivate powerful and engaged research communities.

Acknowledgments

We would like to express our appreciation to Tiebing Liang for DNA preparation and quantification as well as the Social Media Residency with the Mayo Clinic Social Media Network for providing the educational foundation of our digital research approaches. This publication was made possible with support provided to CL from grant numbers KL2TR001106 and UL1TR001108 (A Shekhar, PI) from the National Institutes of Health, National Center for Advancing Translational Sciences, Clinical and Translational Sciences Award.

Conflicts of Interest

None declared.

References

1. Rare Diseases Clinical Research Network. NIH fact sheets URL: https://report.nih.gov/nihfactsheets/ViewFactSheet.aspx?csid=126 [accessed 2017-03-08] [WebCite Cache ID 6onj2bM1c]
2. Grønbæk L, Vilstrup H, Jepsen P. Autoimmune hepatitis in Denmark: incidence, prevalence, prognosis, and causes of death. A nationwide registry-based cohort study. J Hepatol 2014 Mar;60(3):612-617. [doi: 10.1016/j.jhep.2013.10.020] [Medline: 24326217]
3. Manns MP, Czaja AJ, Gorham JD, Krawitt EL, Mieli-Vergani G, Vergani D, et al. Diagnosis and management of autoimmune hepatitis. Hepatology 2010 Jun;51(6):2193-2213. [doi: 10.1002/hep.23584] [Medline: 20513004]
4. European Association for the Study of the Liver. EASL clinical practice guidelines: autoimmune hepatitis. J Hepatol 2015 Dec;63(4):971-1004. [doi: 10.1016/j.jhep.2015.06.030] [Medline: 26341719]
5. Boberg KM, Aadland E, Jahnsen J, Raknerud N, Stiris M, Bell H. Incidence and prevalence of primary biliary cirrhosis, primary sclerosing cholangitis, and autoimmune hepatitis in a Norwegian population. Scand J Gastroenterol 1998 Jan;33(1):99-103. [Medline: 9489916]
6. van Gerven NMF, Verwer BJ, Witte BI, van Erpecum KJ, van Buuren HR, Maijers I, et al. Epidemiology and clinical characteristics of autoimmune hepatitis in the Netherlands. Scand J Gastroenterol 2014 Oct;49(10):1245-1254. [doi: 10.3109/00365521.2014.946083] [Medline: 25123213]
7. Schumacher KR, Stringer KA, Donohue JE, Yu S, Shaver A, Caruthers RL, et al. Social media methods for studying rare diseases. Pediatrics 2014 May;133(5):e1345-e1353 [FREE Full text] [doi: 10.1542/peds.2013-2966] [Medline: 24733869]
8. Franciscus M, Nucci A, Bradley B, Suomalainen H, Greenberg E, Laforte D, et al. Recruitment and retention of participants for an international type 1 diabetes prevention trial: a coordinators' perspective. Clin Trials 2014 Apr;11(2):150-158 [FREE Full text] [doi: 10.1177/1740774513510070] [Medline: 24216218]
9. Lammert C, Comerford M, Love J, Bailey JR. Investigation gone viral: application of the social mediasphere in research. Gastroenterology 2015 Oct;149(4):839-843 [FREE Full text] [doi: 10.1053/j.gastro.2015.08.042] [Medline: 26306768]
10. Zaid T, Burzawa J, Basen-Engquist K, Bodurka DC, Ramondetta LM, Brown J, et al. Use of social media to conduct a cross-sectional epidemiologic and quality of life survey of patients with neuroendocrine carcinoma of the cervix: a feasibility study. Gynecol Oncol 2014 Jan;132(1):149-153 [FREE Full text] [doi: 10.1016/j.ygyno.2013.10.015] [Medline: 24145111]
11. Brustein J. Bloomberg View. Americans now spend more time on Facebook than they do on their pets URL: https://www.bloomberg.com/news/articles/2014-07-23/americans-now-spend-more-time-on-facebook-than-they-do-on-their-pets [accessed 2017-03-08] [WebCite Cache ID 6onjNYnX9]
12. Facebook, Inc. 2017. Facebook reports first quarter 2017 results URL: https://s21.q4cdn.com/399680738/files/doc_financials/2017/Facebook-Reports-First-Quarter-2017-Results.pdf [accessed 2017-05-12] [WebCite Cache ID 6rRSfLULW]
13. Tweet MS, Gulati R, Aase LA, Hayes SN. Spontaneous coronary artery dissection: a disease-specific, social networking community-initiated study. Mayo Clin Proc 2011 Sep;86(9):845-850 [FREE Full text] [doi: 10.4065/mcp.2011.0312] [Medline: 21878595]
14. Lammert C, Nguyen DL, Juran BD, Schlicht E, Larson JJ, Atkinson EJ, et al. Questionnaire based assessment of risk factors for primary biliary cirrhosis. Dig Liver Dis 2013 Jul;45(7):589-594 [FREE Full text] [doi: 10.1016/j.dld.2013.01.028] [Medline: 23490343]
15. Lammert C, Juran BD, Schlicht E, Xie X, Atkinson EJ, de Andrade M, et al. Reduced coffee consumption among individuals with primary sclerosing cholangitis but not primary biliary cirrhosis. Clin Gastroenterol Hepatol 2014 Sep;12(9):1562-1568 [FREE Full text] [doi: 10.1016/j.cgh.2013.12.036] [Medline: 24440215]
16. Hennes EM, Zeniya M, Czaja AJ, Parés A, Dalekos GN, Krawitt EL, et al. Simplified criteria for the diagnosis of autoimmune hepatitis. Hepatology 2008 Jul;48(1):169-176. [doi: 10.1002/hep.22322] [Medline: 18537184]
17. Newton JL, Jones DEJ. Managing systemic symptoms in chronic liver disease. J Hepatol 2012;56 Suppl 1:S46-S55. [doi: 10.1016/S0168-8278(12)60006-3] [Medline: 22300465]
18. Fort D, Wilcox AB, Weng C. Could patient self-reported health data complement EHR for phenotyping? AMIA Annu Symp Proc 2014;2014:1738-1747 [FREE Full text] [Medline: 25954446]

19. Okura Y, Urban LH, Mahoney DW, Jacobsen SJ, Rodeheffer RJ. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. J Clin Epidemiol 2004 Oct;57(10):1096-1103. [doi: 10.1016/j.jclinepi.2004.04.005] [Medline: 15528061]
20. Perrin A. Social media usage 2005-2015. Washington: Pew Internet and American Life Project URL: http://assets.pewresearch.org/wp-content/uploads/sites/14/2015/10/PI_2015-10-08_Social-Networking-Usage-2005-2015_FINAL.pdf [accessed 2018-01-04] [WebCite Cache ID 6wE1Jmhpj]
21. Kulanthaivel A, Fogel R, Jones J, Lammert C. Digital cohorts within the social mediome: an approach to circumvent conventional research challenges? Clin Gastroenterol Hepatol 2017 May;15(5):614-618. [doi: 10.1016/j.cgh.2017.02.015] [Medline: 28433255]
22. Riaz MS, Atreja A. Personalized technologies in chronic gastrointestinal disorders: self-monitoring and remote sensor technologies. Clin Gastroenterol Hepatol 2016 Dec;14(12):1697-1705 [FREE Full text] [doi: 10.1016/j.cgh.2016.05.009] [Medline: 27189911]
23. 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]
24. Bender JL, Jimenez-Marroquin M, Jadad AR. Seeking support on facebook: a content analysis of breast cancer groups. J Med Internet Res 2011;13(1):e16 [FREE Full text] [doi: 10.2196/jmir.1560] [Medline: 21371990]
25. Khatri C, Chapman SJ, Glasbey J, Kelly M, Nepogodiev D, Bhangu A, et al. Social media and internet driven study recruitment: evaluating a new model for promoting collaborator engagement and participation. PLoS One 2015;10(3):e0118899 [FREE Full text] [doi: 10.1371/journal.pone.0118899] [Medline: 25775005]

Abbreviations

AHRN: Autoimmune Hepatitis Research Network
AIH: autoimmune hepatitis
AMA: antimitochondrial antibody
GRACE: Genetic Repository of Autoimmune Liver Diseases and Contributing Exposures
IQR: interquartile range
IU: Indiana University
PBC: primary biliary cholangitis
PSC: primary sclerosing cholangitis
SNS: social networking sites
UDCA: ursodeoxycholic acid

Edited by G Eysenbach; submitted 21.03.17; peer-reviewed by W Davies, A Cyr; comments to author 11.05.17; revised version received 25.06.17; accepted 30.06.17; published 18.01.18.

Please cite as:
Comerford M, Fogel R, Bailey JR, Chilukuri P, Chalasani N, Lammert CS
Leveraging Social Networking Sites for an Autoimmune Hepatitis Genetic Repository: Pilot Study to Evaluate Feasibility
J Med Internet Res 2018;20(1):e14
URL: http://www.jmir.org/2018/1/e14/
doi:10.2196/jmir.7683
PMID:29348111

Viewpoint

From eHealth to iHealth: Transition to Participatory and Personalized Medicine in Mental Health

Sofian Berrouiguet[1,2*], MD, PhD; Mercedes M Perez-Rodriguez[3*], MD, PhD; Mark Larsen[4], D Phil; Enrique Baca-García[5], MD, PhD; Philippe Courtet[6], MD, PhD; Maria Oquendo[7], MD, PhD

[1]Lab-STICC, IMT Atlantique, Université Bretagne Loire, Brest, France
[2]Laboratoire Soins primaires, Santé publique, Registre des cancers de Bretagne Occidentale SPURBO, Equipe d'accueil 7479, Brest, France
[3]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, United States
[4]Black Dog Institute, University of New South Wales, Sydney, Australia
[5]Department of Psychiatry, Fundación Jimenez Diaz Hospital, Autónoma University, Centro de Investigacion en Red Salud Mental, Madrid, Spain
[6]Department of Emergency Psychiatry, University Hospital of Montpellier, University of Montpellier, Montpellier, France
[7]Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States
[*]these authors contributed equally

Corresponding Author:
Sofian Berrouiguet, MD, PhD
Lab-STICC
IMT Atlantique
Université Bretagne Loire
Boulevard Tanguy Prigent, Technopole iroise
Brest,
France
Phone: 33 6 68 20 41 78
Fax: 33 2 98 01 51 57
Email: sofian.berrouiguet@gmail.com

Abstract

Clinical assessment in psychiatry is commonly based on findings from brief, regularly scheduled in-person appointments. Although critically important, this approach reduces assessment to cross-sectional observations that miss essential information about disease course. The mental health provider makes all medical decisions based on this limited information. Thanks to recent technological advances such as mobile phones and other personal devices, electronic health (eHealth) data collection strategies now can provide access to real-time patient self-report data during the interval between visits. Since mobile phones are generally kept on at all times and carried everywhere, they are an ideal platform for the broad implementation of ecological momentary assessment technology. Integration of these tools into medical practice has heralded the eHealth era. Intelligent health (iHealth) further builds on and expands eHealth by adding novel built-in data analysis approaches based on (1) incorporation of new technologies into clinical practice to enhance real-time self-monitoring, (2) extension of assessment to the patient's environment including caregivers, and (3) data processing using data mining to support medical decision making and personalized medicine. This will shift mental health care from a reactive to a proactive and personalized discipline.

(J Med Internet Res 2018;20(1):e2) doi:10.2196/jmir.7412

KEYWORDS

data mining; decision making; mobile phone; Web app; mental health

Introduction

Evolution From eHealth to iHealth

Clinical assessment in psychiatry is usually based on findings from brief, regularly scheduled in-person appointments. Although critically important, this approach reduces assessment to cross-sectional observations that miss essential information about disease course and are subject to recall bias. The mental health provider makes all medical decisions based on this limited information.

After an initial longer assessment, standard follow-up outpatient visits in a psychiatric clinic usually include a mental status examination, a brief update on the history of the patient's present illness (including a safety assessment of risk of self-harm, suicide, and homicide), an assessment of treatment effectiveness and potential side effects, and an updated diagnostic impression and treatment plan. All of this generally happens in under 30 minutes. Except in patients with severe mental illness or disabilities, family members or caregivers rarely attend these visits.

As the duration of a psychiatric outpatient visit becomes increasingly shorter and intervals between appointments longer, it is essential to develop a form of assessment that can more accurately track patients' symptoms between visits [1]. One possible solution is to use personal health records (PHRs), longitudinal health records self-reported by the patient. PHRs can be based on mobile devices (mobile phones, wearable devices) or Web-based self-monitoring. Validity, reliability, and acceptability of this online approach is similar to traditional paper-pencil questionnaires in mental health patients [2]. Surprisingly, although studies have highlighted the value of patient self-reports in clinical assessment, they rarely are routinely implemented [3]. This is despite the fact that many commercial electronic health record (EHR) software packages already allow data entry by patients and caregivers.

Over the last decade, medical assessments have been supported by the increasing use and importance of EHRs that facilitate the portability of pertinent health information across providers and geographic locations. Interinstitutional EHRs further increase efficiency in medical services and provide complete and accurate medical information across providers in different institutions [4]. However, their implementation has only had a modest impact on clinical outcomes and measures of quality of care [5].

Internet features have increased networking possibilities of EHRs, offering new options for patient monitoring. Integration of these tools into medical practice has heralded the electronic health (eHealth) era, integrating new technologies into routine clinical practice. eHealth can also support patient self-monitoring, where both patient and caregivers can update a log of the patient's mental and physical state between medical visits, potentially leading to more accurate assessment. For example, eHealth tools allow an endocrinologist to chart blood glucose levels between visits for a patient with a portable blood glucose meter. Similarly, a mental health clinician could have access to, and be able to chart, changes in sleep, mood, appetite, and other relevant data related to illness course between visits. These data can be collected through Web-based or mobile phone–based self-reports and other assessment tools and sensors. The clinician can take these data into account during the clinical decision-making process before or during the in-person visit.

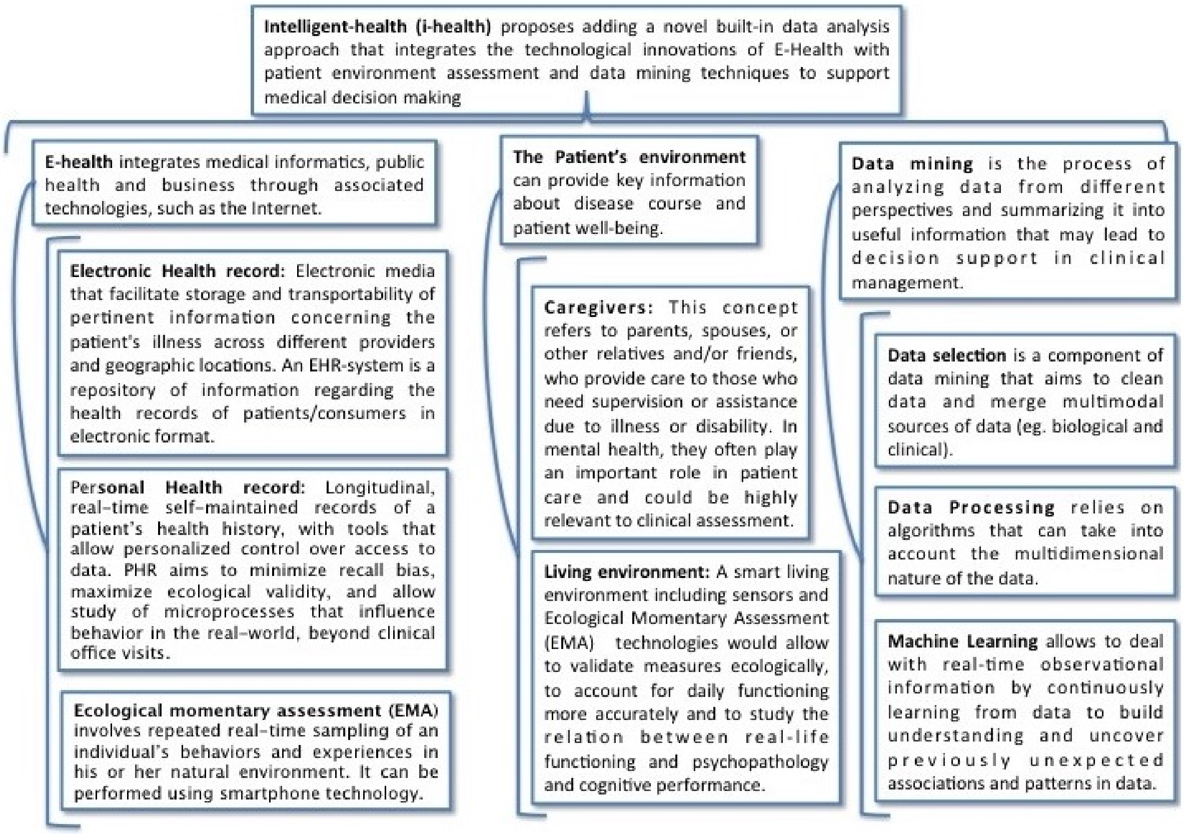
However, we need to go beyond eHealth and move to intelligent health (iHealth). iHealth further builds on and expands eHealth by combining real-time self-monitoring with more contextual information from the patient's environment and novel built-in data analysis tools to enhance medical decision making. The transition from eHealth to iHealth will require integration of comprehensive data from the patient's environment, as reported by the patient or caregiver or captured through sensors in the patient's living environment, and the use of artificial intelligence data-mining techniques to aid clinical decision making and provide more personalized treatment (Figures 1 and 2).

iHealth will allow a mental health provider to receive real-time input from data-mining tools that will help guide clinical decision making. This is particularly important in the field of psychiatry, where the lack of biomarkers and objective biological tests means that most clinical decisions (eg, diagnosis, treatment choice, admission and discharge, risk stratification) are based on signs, symptoms, and behaviors reported to or directly observed by the clinician during the clinical interview. For example, a data-mining iHealth tool may generate a message to alert a clinician that a patient diagnosed with bipolar I disorder is showing a pattern of decreased sleep and increased activity that suggests the imminence of a manic episode. The clinician could decide, based on the alert and other available information, to have a member of the treatment team contact the patient to assess mood stability and determine whether a treatment change is warranted.

In developing countries, iHealth could be used to increase access to specialized care for underserved populations. For example, a data-mining tool could generate data-driven personalized treatment recommendations for a given patient that would be implemented by a general practitioner [6].

Below, we outline a novel iHealth model for clinical assessment and treatment in psychiatry based on (1) incorporation of new technologies into clinical practice to enhance real-time self-monitoring, (2) extension of assessment to the patient's environment, and (3) data processing using data mining to support medical decision making and personalized medicine.

Figure 1. Terms and definition related to e-health.



Intelligent-health (i-health) proposes adding a novel built-in data analysis approach that integrates the technological innovations of E-Health with patient environment assessment and data mining techniques to support medical decision making

E-health integrates medical informatics, public health and business through associated technologies, such as the Internet.

Electronic Health record: Electronic media that facilitate storage and transportability of pertinent information concerning the patient's illness across different providers and geographic locations. An EHR-system is a repository of information regarding the health records of patients/consumers in electronic format.

Personal Health record: Longitudinal, real-time self-maintained records of a patient's health history, with tools that allow personalized control over access to data. PHR aims to minimize recall bias, maximize ecological validity, and allow study of microprocesses that influence behavior in the real-world, beyond clinical office visits.

Ecological momentary assessment (EMA) involves repeated real-time sampling of an individual's behaviors and experiences in his or her natural environment. It can be performed using smartphone technology.

The Patient's environment can provide key information about disease course and patient well-being.

Caregivers: This concept refers to parents, spouses, or other relatives and/or friends, who provide care to those who need supervision or assistance due to illness or disability. In mental health, they often play an important role in patient care and could be highly relevant to clinical assessment.

Living environment: A smart living environment including sensors and Ecological Momentary Assessment (EMA) technologies would allow to validate measures ecologically, to account for daily functioning more accurately and to study the relation between real-life functioning and psychopathology and cognitive performance.

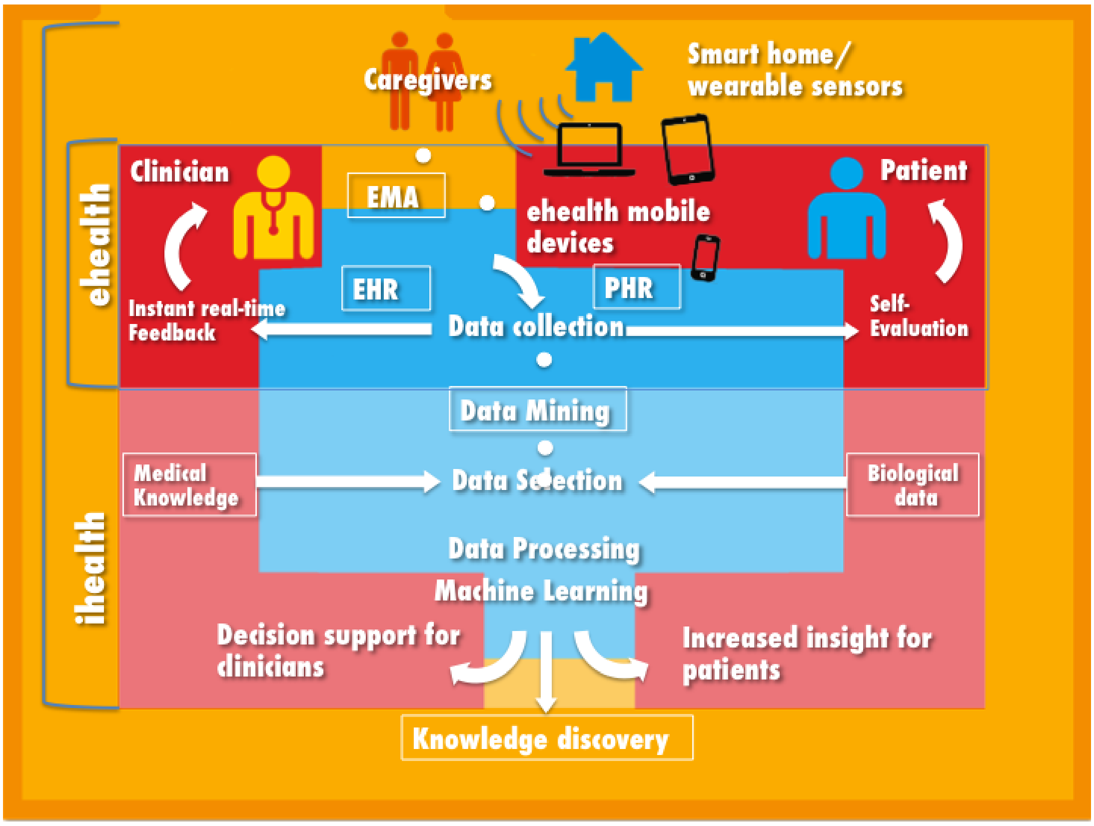
Data mining is the process of analyzing data from different perspectives and summarizing it into useful information that may lead to decision support in clinical management.

Data selection is a component of data mining that aims to clean data and merge multimodal sources of data (eg. biological and clinical).

Data Processing relies on algorithms that can take into account the multidimensional nature of the data.

Machine Learning allows to deal with real-time observational information by continuously learning from data to build understanding and uncover previously unexpected associations and patterns in data.

Figure 2. The transition from electronic health (eHealth) to intelligent health (iHealth). EMA: ecological momentary assessment; EHR: electronic health record; PHR: personal health record.



Emerging Treatment Models

eHealth Era in Mental Health: From Electronic Health Records to Ecological Momentary Assessment

By the end of 2014, there were almost 3 billion Internet users, two-thirds from developing countries, and mobile-broadband subscriptions reached 2.3 billion globally [7]. With such technological advances and reach it is already possible to incorporate Web-based and mobile phone apps into clinical assessment and treatment. Given that psychiatry clinicians have previously relied exclusively on clinical interviews for diagnosis and treatment, the field could deeply benefit from this new source of data collected in real time covering information about the patient's health state between visits. Mobile phones are generally kept on at all times and carried everywhere, making them an ideal platform for the broad implementation of ecological momentary assessment (EMA) technology. EMA involves repeated sampling of subjects' behaviors and experiences in real time, in their natural environment. Patient self-monitoring can rely on EMA procedures and lead to participatory medicine. EMA has been successfully used for real-time self-reporting of symptoms and behavior—for example, Husky et al [8] showed the utility and feasibility of using EMAs to study suicidal ideation [8].

The emergence of smart homes and the development of sensor technologies allows the nonintrusive collection of activity data [9], allowing objective analysis of an individual's behavioral patterns. Thus, health-related events such as activities of daily living (ADLs, feeding, taking care of one's personal hygiene, dressing) can be captured without the patient's active participation. Monitoring behavioral patterns of psychiatric patients and their ability to carry out their ADLs in their living environment will likely improve knowledge about disease course. The detection of changes in patterns of behavior may help to detect emerging disorders [10] and study the relationship between functioning and cognitive performance or illness course [11]. In mental health, this approach is still at an early stage but has already shown promising results in the monitoring of depressive symptoms in cognitively impaired patients [12] and of activity in patients with schizophrenia [13]. Smart Home and Ambient Assisted Living (SHAAL) systems use sensors and other devices that are either wearable or integrated in the patient's home and have been used to assess the effect of negative symptoms and cognitive impairment on ADL functions [14] and detect emerging disorders based on changes in the patient's behavior [15].

Future studies need to assess the impact of the support provided by these smart home devices on patient outcomes and the sensitivity and specificity of the data collection devices. The cost-effectiveness of these new monitoring approaches also needs to be assessed. Furthermore, the ethical and legal aspects have to be addressed, taking into account privacy and medical confidentiality issues [16].

Including Patient Environment in Mental Health Monitoring

Many psychiatric disorders are chronic illnesses associated with high levels of disability [17], making it a challenge for some psychiatric patients to live independently. Therefore, caregivers often play a critical role in the lives of those with serious mental illness [18].

Studies have shown that caregivers and close contacts are reliable sources of information about patients with psychiatric disorders [19]. Traditional psychiatric assessment, however, does not always include information from caregivers due to time constraints and concerns about confidentiality [20]. By excluding caregivers from assessments, clinicians may miss an opportunity to obtain additional valuable information about the illness course [21].

The new technologies described above, including PHRs and EMA, can easily include not only the patient's own reports and activity but also those from a caregiver [22]. Capturing caregivers' reports through an EMA approach could provide a more accurate assessment of the illness course in a given patient. Both the patient and the caregiver would be able to enter updates in an online log of the patient's mental and physical state between outpatient visits.

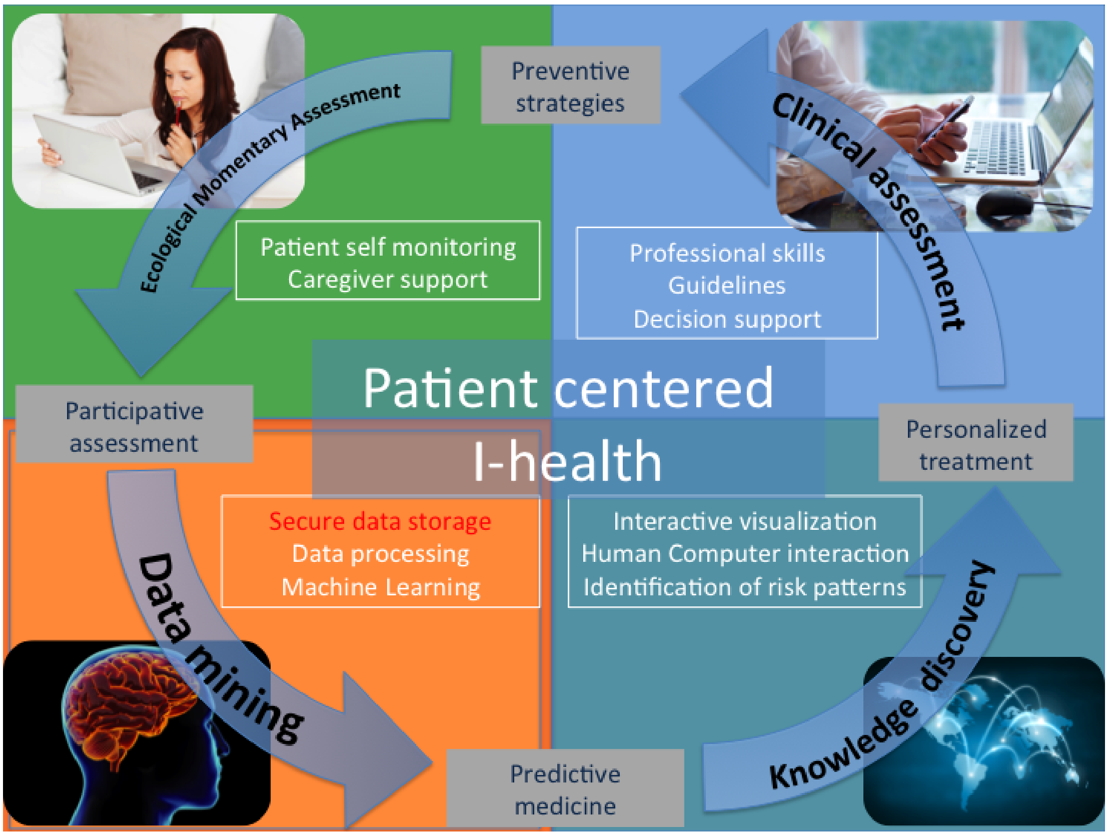
Involving caregivers in the clinical assessment may also help to decrease the caregiver burden, including physical and psychological stress, social pressure, and financial costs associated with care giving. Unfortunately, studies focusing on caregiver burden are scarce and have used small heterogeneous patient samples [23]. Identifying and modifying caregivers' burdens through EMA might help to reduce the level of burden and its negative effects, both on the caregivers and on patient outcomes [20].

We are advancing toward a double paradigm shift. First, the integration of patient and caregiver data through EMA and PHRs into routine psychiatric assessment will give clinicians access to real-time reports and behaviors of patients and caregivers. Second, as described in the next section, the application of data-mining techniques to the EMA and PHR data will support and enhance medical decision making. These data can be analyzed using data mining tools in order to develop predictive models and personalized treatments. This will shift mental health from a reactive to a proactive discipline, leading to decision support systems for clinicians, as summarized in Figure 1.

Transforming Data into Knowledge Through iHealth

The traditional method of turning data into knowledge has relied on manual data analysis and interpretation of results in order to find useful patterns to support decision making. The enormous amount and complexity of the ecological data that can be obtained through EMA and PHRs [24,25] make manual analysis by end users (that is, clinicians) difficult or impossible. Making sense of enormous datasets is a common challenge in the Big Data era that is best overcome using data mining techniques.

Figure 3. iHealth in the context of the 4P (predictive, personalized, preventive and participatory) model of mental health.



Data mining is a set of techniques that can be used to explore treatment and outcome questions in large clinical databases and help develop algorithms and guidelines for problems where controlled data are difficult to obtain. The data mining process includes several steps, including data selection, data processing, and machine learning (described below). Data mining techniques can be used to find relationships and patterns between EMA and PHR data and neurobiological data. As an example, experiments have shown that connections between momentary mental states and environments are sensitive to genetic effects, emphasizing gene-environment interaction [26]. This may lead to advances in clinical decision making, incorporating clinical, ecological, and biological data [27,28].

Machine learning techniques seek to answer the question: How can we build computer systems that automatically improve with experience, and what are the fundamental laws that govern all learning processes? Machine learning techniques allow processing of real-time observational information by continuously learning from data to build understanding and uncover previously unexpected associations and patterns in data [29]. Predictive and explanatory models might use individual patient data to predict future events like the probability of a patient attempting suicide in a given time interval. Machine learning models are continuously updated to refine and improve their clinical applicability [30]. This process may significantly improve decision making [31] and knowledge discovery (Figures 2 and 3). Specifically, data mining allows for the exploration of risk factors, patterns of symptom evolution, and identification of high-risk subgroups [32].

Discussion

iHealth as a Practical and Conceptual Challenge for Mental Health Professionals and Patients

Despite its many potential advantages over traditionally used psychiatric assessment tools, iHealth still faces several risks and challenges related to human factors. First, electronic monitoring devices are often experienced by physicians as financial, technical, or time-consuming threats. Furthermore, psychological barriers are also important to consider [33]. With the implementation of iHealth tools, physicians may be concerned about the loss of control over patient information and the decision-making process, since these data can be shared with and assessed by others. It is uncertain if clinicians will be willing to share the decision-making process with machine-learning tools and a data-analysis team.

It could be argued that eHealth and iHealth may be most useful for severe and disabling psychiatric disorders, which are characterized by poor insight and often require significant involvement of caregivers and family members in the clinical care. Some examples could be major neurocognitive disorders (dementia), severe psychotic disorders (schizophrenia), or autism spectrum disorders. However, it should be noted that growing evidence suggests that family involvement and knowledge about the illness and regular contact between patients and caregivers improve outcomes across psychiatric disorders [34-36].

Another limitation of iHealth is that it has been argued that some psychiatric conditions are not amenable to self-reporting [37]. If a patient is in crisis or suffers from cognitive impairment or psychosis, EMA or PHR assessments may be of limited utility. In such situations, however, sensor-based technologies can still provide valuable objective data [6].

It may also be argued that requiring the patient to constantly keep track of their mental and physical state puts more responsibility and burden on them. Whether this would have positive or negative effects remains unexplored [38]. eHealth and iHealth interventions require an important time commitment from the participants, particularly for those who collect regular daily EMA data. This approach may be significantly more invasive than asking a participant to complete a retrospective questionnaire or answer a question through a traditional interview. The risk of intrusiveness into daily lives exists, but this issue was not assessed in the articles we reviewed or in other reviews in the field. While a growing number of health care systems in developed countries use some form of EHR or PHR, very few use eHealth for the delivery of mental health care. Therefore, increasing the update of eHealth interventions is a crucial step toward realizing the potential of iHealth.

Data privacy concerns may deter patients from sharing personal data related to mental health, but data security procedures have been a routine part of data mining from the outset [39]. Human factors including acceptability aspects regarding the technology and control over personal data will be critical to accomplish the transition to iHealth.

Conclusion

Despite the challenges and limitations iHealth may lead to improved clinically integrated decision-making tools and personalized medicine practice, tailoring medical treatment to each individual patient. Building on the advances in mental health assessment brought about by eHealth, iHealth will provide personalized clinical information outside of clinical visits and integration of real-time multimodal patient and caregiver data using data mining technologies. This will allow for more precise and effective clinical assessment and decision making. The ability to mine large databases for new hypotheses regarding clinical and environmental dynamic patterns of psychiatric illness through iHealth could also change clinical practice. These possibilities serve key public health needs and offer intriguing and novel opportunities for collaboration, knowledge production, and data analysis.

Acknowledgements

This research was partially support by Instituto de Salud Carlos III (PI16/01852 Grant), Plan Nacional de Drogas (20151073 Project), and American Foundation for Suicide Prevention (LSRG-1-005-16). SB’s work was supported by Fondation de l’Avenir, the French Embassy in Madrid; MMPR's work was supported by a National Alliance for Research on Schizophrenia and Depression (NARSAD) Young Investigator Award (YIA) grant and a KL2 Faculty Scholar (KL2TR001435) grant (PI: Perez-Rodriguez).

Conflicts of Interest

None declared.

References

1. Olfson M, Kroenke K, Wang S, Blanco C. Trends in office-based mental health care provided by psychiatrists and primary care physicians. J Clin Psychiatry 2014 Mar;75(3):247-253. [doi: 10.4088/JCP.13m08834] [Medline: 24717378]
2. Zimmerman M, Martinez JH. Web-based assessment of depression in patients treated in clinical practice: reliability, validity, and patient acceptance. J Clin Psychiatry 2012 Mar;73(3):333-338. [doi: 10.4088/JCP.10m06519] [Medline: 22154899]
3. Sartorius N. Patient-reported outcomes in psychiatry. Dialogues Clin Neurosci 2014 Jun;16(2):123-124 [FREE Full text] [Medline: 25152651]
4. Blecker S, Goldfeld K, Park N, Shine D, Austrian J, Braithwaite R, et al. Electronic health record use, intensity of hospital care, and patient outcomes. Am J Med 2014 Mar;127(3):216-221 [FREE Full text] [doi: 10.1016/j.amjmed.2013.11.010] [Medline: 24333204]
5. Windle JR, Windle TA. Electronic health records and the quest to achieve the Triple Aim. J Am Coll Cardiol 2015 May 12;65(18):1973-1975 [FREE Full text] [doi: 10.1016/j.jacc.2015.03.038] [Medline: 25953749]
6. Medhanyie AA, Spigt M, Yebyo H, Little A, Tadesse K, Dinant G, et al. Quality of routine health data collected by health workers using smartphone at primary health care in Ethiopia. Int J Med Inform 2017 May;101:9-14. [doi: 10.1016/j.ijmedinf.2017.01.016] [Medline: 28347452]
7. The World in 2014. ICT facts and Figures. 2014. URL: http://www.itu.int/en/ITU-D/Statistics/Pages/facts/default.aspx [accessed 2017-12-07] [WebCite Cache ID 6vWsb6zpY]
8. Husky M, Olié E, Guillaume S, Genty C, Swendsen J, Courtet P. Feasibility and validity of ecological momentary assessment in the investigation of suicide risk. Psychiatry Res 2014 Dec 15;220(1-2):564-570. [doi: 10.1016/j.psychres.2014.08.019] [Medline: 25155939]
9. Soulas J, Lenca P, Thépaut A. Unsupervised discovery of activities of daily living characterized by their periodicity and variability. Eng Appl Artif Intel 2015:90-102.
10. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. J Med Internet Res 2015;17(7):e175 [FREE Full text] [doi: 10.2196/jmir.4273] [Medline: 26180009]

11. Tang P, Venables T. 'Smart' homes and telecare for independent living. J Telemed Telecare 2000;6(1):8-14. [doi: 10.1258/1357633001933871] [Medline: 10824384]
12. Galambos C, Skubic M, Wang S, Rantz M. Management of dementia and depression utilizing in-home passive sensor data. Gerontechnology 2013 Jan 01;11(3):457-468 [FREE Full text] [Medline: 24049513]
13. Stip E, Rialle V. Environmental cognitive remediation in schizophrenia: ethical implications of smart home technology. Can J Psychiatry 2005 Apr;50(5):281-291. [doi: 10.1177/070674370505000509] [Medline: 15968845]
14. Bowie CR, Reichenberg A, Patterson TL, Heaton RK, Harvey PD. Determinants of real-world functional performance in schizophrenia subjects: correlations with cognition, functional capacity, and symptoms. Am J Psychiatry 2006 Mar;163(3):418-425. [doi: 10.1176/appi.ajp.163.3.418] [Medline: 16513862]
15. Semkovska M, Bédard M, Godbout L, Limoge F, Stip E. Assessment of executive dysfunction during activities of daily living in schizophrenia. Schizophr Res 2004 Aug 01;69(2-3):289-300. [Medline: 15469200]
16. Sorell T, Draper H. Telecare, surveillance, and the welfare state. Am J Bioeth 2012;12(9):36-44. [doi: 10.1080/15265161.2012.699137] [Medline: 22881854]
17. Chakrabarti S, Kulhara P, Verma SK. Extent and determinants of burden among families of patients with affective disorders. Acta Psychiatr Scand 1992 Sep;86(3):247-252. [Medline: 1414422]
18. González N, Padierna A, Martín J, Aguirre U, Quintana JM. Predictors of change in perceived burden among caregivers of patients with eating disorders. J Affect Disord 2012 Aug;139(3):273-282. [doi: 10.1016/j.jad.2012.02.013] [Medline: 22391516]
19. Berk L, Berk M, Dodd S, Kelly C, Cvetkovski S, Jorm AF. Evaluation of the acceptability and usefulness of an information website for caregivers of people with bipolar disorder. BMC Med 2013 Jul 11;11:162 [FREE Full text] [doi: 10.1186/1741-7015-11-162] [Medline: 23844755]
20. 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]
21. Teferra S, Hanlon C, Beyero T, Jacobsson L, Shibre T. Perspectives on reasons for non-adherence to medication in persons with schizophrenia in Ethiopia: a qualitative study of patients, caregivers and health workers. BMC Psychiatry 2013 Jun 17;13:168 [FREE Full text] [doi: 10.1186/1471-244X-13-168] [Medline: 23773362]
22. Runyan JD, Steenbergh TA, Bainbridge C, Daugherty DA, Oke L, Fry BN. A smartphone ecological momentary assessment/intervention app for collecting real-time data and promoting self-awareness. PLoS One 2013;8(8):e71325 [FREE Full text] [doi: 10.1371/journal.pone.0071325] [Medline: 23977016]
23. Caqueo-Urízar A, Miranda-Castillo C, Lemos GS, Lee MS, Ramírez PM, Mascayano TF. An updated review on burden on caregivers of schizophrenia patients. Psicothema 2014 May;26(2):235-243. [doi: 10.7334/psicothema2013.86] [Medline: 24755026]
24. 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]
25. Akil H, Martone ME, Van EDC. Challenges and opportunities in mining neuroscience data. Science 2011 Feb 11;331(6018):708-712 [FREE Full text] [doi: 10.1126/science.1199305] [Medline: 21311009]
26. Dominguez M, Wichers M, Lieb R, Wittchen H, van Os J. Evidence that onset of clinical psychosis is an outcome of progressively more persistent subclinical psychotic experiences: an 8-Year cohort study. Schizophrenia Bull 2010;37(1):84-93.
27. Jin S, Tan R, Jiang Q, Xu L, Peng J, Wang Y, et al. A generalized topological entropy for analyzing the complexity of DNA sequences. PLoS One 2014;9(2):e88519 [FREE Full text] [doi: 10.1371/journal.pone.0088519] [Medline: 24533097]
28. Sultan M, Wigle DA, Cumbaa CA, Maziarz M, Glasgow J, Tsao MS, et al. Binary tree-structured vector quantization approach to clustering and visualizing microarray data. Bioinformatics 2002;18 Suppl 1:S111-S119. [Medline: 12169538]
29. Oquendo MA, Baca-Garcia E, Artés-Rodríguez A, Perez-Cruz F, Galfalvy HC, Blasco-Fontecilla H, et al. Machine learning and data mining: strategies for hypothesis generation. Mol Psychiatry 2012 Oct;17(10):956-959. [doi: 10.1038/mp.2011.173] [Medline: 22230882]
30. Markowetz A, Baszkiewicz K, Montag C, Switala C, Schlaepfer T. Psycho-informatics: Big Data shaping modern psychometrics. Med Hypotheses 2014:405-411.
31. Lee D. Decision making: from neuroscience to psychiatry. Neuron 2013 Apr 24;78(2):233-248 [FREE Full text] [doi: 10.1016/j.neuron.2013.04.008] [Medline: 23622061]
32. Baca-García E, Perez-Rodriguez MM, Basurte-Villamor I, Saiz-Ruiz J, Leiva-Murillo JM, Santiago-Mozos R, et al. Using data mining to explore complex clinical decisions: a study of hospitalization after a suicide attempt. J Clin Psychiatry 2006 Jul;67(7):1124-1132. [Medline: 16889457]
33. Boonstra A, Broekhuis M. Barriers to the acceptance of electronic medical records by physicians from systematic review to taxonomy and interventions. BMC Health Serv Res 2010;10:231 [FREE Full text] [doi: 10.1186/1472-6963-10-231] [Medline: 20691097]
34. Jolley S, Ferner H, Bebbington P, Garety P, Dunn G, Freeman D, et al. Delusional belief flexibility and informal caregiving relationships in psychosis: a potential cognitive route for the protective effect of social support. Epidemiol Psychiatr Sci 2014 Dec;23(4):389-397. [doi: 10.1017/S2045796013000553] [Medline: 24160466]

35. Revier CJ, Reininghaus U, Dutta R, Fearon P, Murray RM, Doody GA, et al. Ten-year outcomes of first-episode psychoses in the MRC SOP-10 study. J Nerv Ment Dis 2015 May;203(5):379-386 [FREE Full text] [doi: 10.1097/NMD.0000000000000295] [Medline: 25900547]
36. Scott J, Colom F, Pope M, Reinares M, Vieta E. The prognostic role of perceived criticism, medication adherence and family knowledge in bipolar disorders. J Affect Disord 2012 Dec 15;142(1-3):72-76. [doi: 10.1016/j.jad.2012.04.005] [Medline: 22944191]
37. Bauer M, Glenn T, Grof P, Pfennig A, Rasgon N, Marsh W. Self-reported data from patients with bipolar disorder: frequency of brief depression. J Affect Disord 2007;101:233.
38. 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]
39. Regola N, Chawla NV. Storing and using health data in a virtual private cloud. J Med Internet Res 2013 Mar 13;15(3):e63 [FREE Full text] [doi: 10.2196/jmir.2076] [Medline: 23485880]

Abbreviations

ADL: activity of daily living
eHealth: electronic health
EHR: electronic health record
EMA: ecological momentary assessment
iHealth: intelligent health
PHR: personal health record
SHAAL: Smart Home and Ambient Assisted Living

Edited by G Eysenbach; submitted 11.04.17; peer-reviewed by N Bragazzi, J Torous, K Gurditta, HY Huang, TR Soron, M Pompili, R Shrestha; comments to author 17.06.17; revised version received 08.08.17; accepted 13.09.17; published 03.01.18.

Please cite as:
Berrouiguet S, Perez-Rodriguez MM, Larsen M, Baca-García E, Courtet P, Oquendo M
From eHealth to iHealth: Transition to Participatory and Personalized Medicine in Mental Health
J Med Internet Res 2018;20(1):e2
URL: http://www.jmir.org/2018/1/e2/
doi:10.2196/jmir.7412
PMID:29298748

Original Paper

Gender Differences in and the Relationships Between Social Anxiety and Problematic Internet Use: Canonical Analysis

Mustafa Baloğlu[1], PhD; Hatice İrem Özteke Kozan[2], PhD; Şahin Kesici[2], PhD

[1]Department of Special Education, Faculty of Education, Hacettepe University, Ankara, Turkey

[2]Department of Counseling, Ahmet Kelesoglu Faculty of Education, Necmettin Erbakan University, Konya, Turkey

Corresponding Author:
Mustafa Baloğlu, PhD
Department of Special Education
Faculty of Education
Hacettepe University
Beytepe Campus
Ankara, 06610
Turkey
Phone: 90 3127806315
Fax: 90 3127806314
Email: baloglu@hotmail.com

Abstract

Background: The cognitive-behavioral model of problematic Internet use (PIU) proposes that psychological well-being is associated with specific thoughts and behaviors on the Internet. Hence, there is growing concern that PIU is associated with psychological impairments.

Objective: Given the proposal of gender schema theory and social role theory, men and women are predisposed to experience social anxiety and engage in Internet use differently. Thus, an investigation of gender differences in these areas is warranted. According to the cognitive-behavioral model of PIU, social anxiety is associated with specific cognitions and behaviors on the Internet. Thus, an investigation of the association between social anxiety and PIU is essential. In addition, research that takes into account the multidimensional nature of social anxiety and PIU is lacking. Therefore, this study aimed to explore multivariate gender differences in and the relationships between social anxiety and PIU.

Methods: Participants included 505 college students, of whom 241 (47.7%) were women and 264 (52.3%) were men. Participants' ages ranged from 18 to 22 years, with a mean age of 20.34 (SD=1.16). The Social Anxiety Scale and Problematic Internet Use Scale were used in data collection. Multivariate analysis of variance (MANOVA) and canonical correlation analysis were used.

Results: Mean differences between men and women were not statistically significant in social anxiety (λ=.02, F3,501=2.47, *P*=.06). In all three PIU dimensions, men scored higher than women, and MANOVA shows that multivariate difference was statistically significant (λ=.94, F3,501=10.69, *P*<.001). Of the canonical correlation functions computed for men, only the first was significant (Rc=.43, λ=.78, χ29=64.7, *P*<.001) and accounted for 19% of the overlapping variance. Similarly, only the first canonical function was significant for women (Rc=.36, λ=.87, χ29=33.9, *P*<.001), which accounted for 13% of the overlapping variance.

Conclusions: On the basis of the findings, we conclude that enhanced educational opportunities for women and their increasing role in the society have led women to become more active and thus closed the gap in social anxiety levels between men and women. We found that men showed more difficulties than women in terms of running away from personal problems (ie, social benefit), used the Internet more excessively, and experienced more interpersonal problems with significant others due to Internet use. We conclude that men are under a greater risk of social impairments due to PIU. Our overall conclusion is that there is a substantial amount of association between social anxiety and PIU and the association is stronger for men than it is for women. We advise that future research continue to investigate PIU and social anxiety as multidimensional constructs.

(J Med Internet Res 2018;20(1):e33) doi:10.2196/jmir.8947

KEYWORDS

anxiety; Internet; sex characteristics; social anxiety disorder; addictive behavior

Introduction

Adam is a 20-year-old college student who loiters a few minutes in front of his college counselor's office and rehearses impatiently what he is to say to her. He decides that he will quickly list all his problems once and for all:

> *I have serious problems and I do not know what to do with them. I've found a world on the Internet and I do not let anyone enter into that world. To be honest with you, my mom bugs me to go out more; however, I don't have the guts! I am constantly reminded of the pitfalls of talking to people. The comfort of the Internet is very soothing. I do not know why I am avoiding people. I guess I fear that the more I open up to others, the more I give them to criticize me. I mean everything about me…the way I dress, talk, act…everything! I cannot live with such criticism! Whenever I am being criticized I feel worthless. I am not good at ice breakers, you know. But the Internet is not like that. Nobody really knows who you are on the Internet. I can say anything I wanna and nobody knows who you are! I can post a handsome picture on my profile and have more time to think before I say what I wanna to say. I get more 'likes' and feel accepted. I do not think anything is wrong with that! But my mom says I have problems and I've got to see a shrink. What do you think?*

While he was rehearsing, the door opens and the counselor invites him in. All he was thinking was suddenly gone out of his mind and he starts sweating and shaking. He was not even able to look her in the eye…

The use of the Internet may result in healthy (ie, positive) or unhealthy (ie, negative) consequences [1]. Internet use can be identified with positive behavioral changes, psychological comfort, and the ease of accessing the available material. On the other hand, problematic Internet use (PIU) may lead to academic, social, occupational, or psychological impairments [2]. The distinction between positive and negative use of the Internet primarily depends on the amount of time spent and the types of activities engaged on it [3].

Problematic Internet Use

PIU is a comprehensive term that encompasses various unhealthy consequences of the Internet, including negative use, social benefits, excessive use, addiction, social comfort, depression, impulse control, and distraction [4]. Beard and Wolf [5] assert that PIU is a more encompassing construct compared with the term Internet addiction, even though the two share a few common characteristics and may be used interchangeably in the literature. They argue that PIU does not necessarily contain some of the symptoms that are available in Internet addiction. We opt to study PIU in our research.

The cognitive-behavioral model of PIU proposes that psychological well-being is associated with specific cognitions and behaviors on the Internet [4,6,7]. Hence, there is growing concern that PIU is associated with psychological impairments [2,8]. Numerous measurement instruments have been developed to objectively assess PIU (eg, [9,10]). The detrimental effects of PIU on psychological well-being are documented in the literature (for a review see [1]).

According to the displacement hypothesis, time spent on the Internet takes away time spent in real-life interactions [11]. Thus, psychological problems may occur when one engages more in Web-based relationships than real-life interactions [12]. Furthermore, *the lonely drawn to the Internet* hypothesis proposes that individuals who feel lonely and are socially more anxious engage in Web-based activities as a solution [13]. Studies on loneliness, shyness, and social anxiety show that Web-based communication may be sought as a potential escape (eg, [14]). In addition, insecurity is also found to have significant effects in the development of PIU (eg, [15]).

Socially anxious individuals seek environments in which they can engage with others more conveniently. The Internet is an available ground for such people to disclose their inner selves, interact with others, and gratify their needs . However, Internet use may become problematic for individuals in different ways. A review of the relevant literature based on gender schema theory and social role theory shows that men and women experience social anxiety and PIU differently. Therefore, this study aimed to report multivariate gender differences in social anxiety and PIU. In addition, we explored the interaction between social anxiety and PIU across men and women.

Social Anxiety

According to the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-5*) [16], individuals who have difficulties in forming and maintaining relationships may reveal anxious attitudes in "social relationships." In this respect, *social anxiety* refers to a strong fear of being judged by others and the corresponding feelings of shame [17]. Social anxiety has been the subject of numerous studies in recent years [18]. Studies usually show that individuals with intense fear of being evaluated by others and those who worry extensively about being criticized in social situations prefer environments in which they can reveal themselves more comfortably [19]. The lonely drawn to the Internet hypothesis proposes that one such medium is the *Internet*. Those who experience difficulties in face-to-face communication prefer Web-based relationships [20]. Therefore, the Internet is an important source in coping with social anxiety, and it provides opportunities for social communication [21]. Additionally, technology-based applications are effectively used in the assessment [22] and treatment [23] of social anxiety.

Studies on social anxiety and Internet use show that individuals facing considerable difficulty in social interactions tend to use the Internet more frequently [15]. Given the proposal of the cognitive-behavioral model of PIU and the existence of significant associations between psychological health and PIU in previous studies, an investigation of the relationships between social anxiety and PIU is warranted.

Gender differences are partially determined by sociocognitive factors [24] such as social roles (ie, [25]) and gender roles (ie, [26,27]). According to gender role theory [28,29], different attributes are attached to femininity and masculinity in the culture.

Table 1. A summary of the problematic Internet use literature.

Study	Participants	Analysis	Major finding
Cao et al [43]	17,599 Chinese adolescents	*t* test	Males were higher than females in PIU[a]
Hetzel-Riggin and Pritchard [44]	425 American undergraduates	Hierarchical multiple regression	Social anxiety was a predictor for men's PIU, but depression was a predictor for women's PIU
Kormas et al [45]	866 Greek adolescents	Descriptive statistics	Maladaptive Internet users are more likely to be males
Mottram and Fleming [46]	272 Australian undergraduates	Multivariate analysis of variance	Males report more problems related to Internet use than females
Durkee et al [47]	11,956 adolescents in different European countries	Descriptive statistics	Female students were higher in maladaptive Internet use, whereas males were higher in pathological Internet use
Schimmenti et al [48]	310 Italian high school students	Descriptive statistics	Males showed more risky behaviors in PIU

[a] PIU: problematic Internet use.

For example, athletic, competitive, dominant, and aggressive are a few of the attributes of masculinity, whereas affectionate, sensitive, warm, and sympathetic are some of the common attributes of femininity in many cultures. Gender role theory has eventually evolved to gender schema theory [30,31], which purports that individuals develop gender-appropriate cognitive schemas early in childhood through social learning to behave consistent to their biological gender [32]. In addition, social role theory considers gender schemas as dynamic structures of the culture and hypothesizes that individuals display different social roles depending upon their societal and cultural surroundings. For example, we tend to behave quite differently when we are among family members from when we are among colleagues. Social role theory indicates that the expectations of the culture from men and women reveal themselves in the use of power and statute [33] and *gender role beliefs* are shaped by the perceptions of the social expectations [34]. Therefore, according to Bussey and Bandura [24], gender-relevant outcomes are social rather than instinctual. In sum, cognitive gender schemas largely affect individuals' thought processes and behaviors (eg, [26,29,31]) and eventually gender plays a significant role in explaining differences in social anxiety as well as PIU. Therefore, gender differences in psychological health (eg, [35]) and Internet use (eg, [36]) have long been investigated.

As we summarized above, cognitive and behavioral constructs prime men and women to think, feel, believe, or act differently. Anxiety is one such emotion that men and women experience differently. Spielberger [37] conceptualized anxiety as trait anxiety and state anxiety; in which the former refers to stable individual differences in anxiety proneness, whereas the later refers to a content-dependent, transitory emotional condition. Although studies are not in complete consensus, research largely indicates both in general anxiety [35] and various types of state anxieties, women score higher than men (eg, [38-40]). Social anxiety can also be regarded as a specific type of state anxiety. Even though a few studies found the contrary (eg, [41]), most epidemiological and community-based investigations have found that the prevalence of social anxiety is higher among women [42]. Therefore, we hypothesized that women would score higher on social anxiety than men.

Gender differences have also been studied in relation to the Internet. Studies largely show that men are more likely to use the Internet and encounter higher levels of PIU than women (see Table 1). Therefore, we also hypothesized that men would score higher on PIU than women.

In sum, given the proposal of gender schema theory and social role theory, men and women are predisposed to experience social anxiety and use the Internet differently. Therefore, an investigation of gender differences in these areas is warranted. According to the cognitive-behavioral model of PIU, social anxiety is associated with specific cognitions and behaviors on the Internet. Therefore, an investigation of the association between social anxiety and PIU is essential. In addition, research that takes into account the multidimensional nature of both social anxiety and PIU is lacking. Therefore, we developed and tested the following research hypotheses.

Hypotheses

Bem [30,31] theorized that individuals develop cognitive schemas throughout childhood to display gender-appropriate behaviors. Such cognitive schemas eventually predispose men and women to reason and behave differently. For example, it is well documented that women place more value on the views of others regarding their own appearances [49]. Thus, it can be hypothesized that women would suffer more from social anxiety. Similarly, researchers show that men and women behave differently on the Internet. Durkee et al [47] found that male adolescents preferred Web-based games, whereas female adolescents preferred social interactions such as chat rooms on the Internet. In addition, Weiser [50] found that men use the Internet more for entertainment purposes, whereas women use it more for interpersonal communication. On the basis of such previous findings, we hypothesized that women would score significantly higher than men in social anxiety components (hypothesis 1), but men would score significantly higher than women in PIU (hypothesis 2).

According to the displacement hypothesis, because socially anxious individuals spend less time with others in real life, they will tend to dwell more on the Internet, leading to more serious consequences of PIU. Supporting these theories, Fehm et al [19] conclude that people who experience negative feelings tend to prefer safer environments such as the Internet. In addition, it can be postulated that the Internet provides a secure environment for individuals who display the symptoms of social anxiety. Similarly, Weinstein et al [18] found statistically significant associations between social anxiety and Internet addiction in two samples of young adults (r=.41 and r=.34, P<.001). Chiang and Hsiao [20] indicated that individuals who prefer Web-based communication tend to also show more PIU. However, only a few studies investigated the relationship between PIU and psychological well-being separately for men and women. Huang [1] advises that the relationship between PIU and psychological well-being should be explored across men and women. Therefore, we hypothesized that there would be a positive relationship between social anxiety dimensions and PIU dimensions both for men and women (hypothesis 3).

Methods

Participants

Because social anxiety is hypothesized to manifest from midteen years [16] and because PIU is more prevalent among college students [51], we decided to include a group of college students in our sample. A total of 505 college students who were enrolled in classes in two large state universities in Turkey voluntarily participated in the study. Of the group, 47.7% (241/505) were women and 52.3% (264/505) were men. Participants' ages ranged from 18 to 22 years, with a mean age of 20.34 years (SD=1.16). None of the participants were clinically diagnosed in terms of either social anxiety or PIU. The number of hours of Internet use among participants ranged from 1 to 105 hours per week with a mean of 14.84 hours (SD 15.94).

Measures

The Social Anxiety Scale (SAS) and the Problematic Internet Use Scale (PIUS) were used to collect the data. The SAS was developed by Özbay and Palancı [52] to assess social anxiety within the frame of social avoidance, criticism anxiety, and self-depreciation. The scale includes 30 Likert-type items, and higher scores on the total and subscales indicate higher levels of social anxiety, social avoidance, criticism anxiety, and self-depreciation. The psychometric properties of the SAS were found adequate [52]. Internal consistency of the SAS and its subscales were found acceptable in the current sample for both men and women (Table 2).

The PIUS was developed by Ceyhan et al [53] as a 33-item, 5-point Likert-type self-report instrument and includes the following three subscales: negative use, social benefit, and excessive use. Higher scores on the scale indicate higher levels of PIU. The scale was found to have adequate validity and reliability properties [53]. Internal consistency of the PIUS and its subscales was found to be adequate in the current sample for both men and women (Table 2).

Procedure

Human Subjects Committee had reviewed and approved the study protocol before we started the data collection procedure. Students were recruited in their classes and informed about the purposes of the study. After signing the consent form, the counterbalanced research packets were administered during class hours, within approximately 30 min. Data were screened for the assumptions of parametric statistics. Multivariate outliers, normality, and homogeneity of variances were tested. Gender differences on the subscales of the SAS and the PIUS were investigated using multivariate analysis of variance (MANOVA). In addition, two separate canonical correlation analyses were conducted to investigate how a set of social anxiety variables was related to a set of PIU variables independently for both men and women.

Table 2. Descriptive statistics for the Social Anxiety Scale (SAS) and the Problematic Internet Use Scale (PIUS) for men and women.

Variables[a,b]	SAS	Social Avoidance	Criticism Anxiety	Self-Deprecation	PIUS	Excessive Use	Social Benefit	Negative Use
Social Avoidance	.94 (.94)							
Criticism Anxiety	.90 (.93)	.76 (.80)						
Self-Deprecation	.87 (.90)	.72 (.76)	.68 (.78)					
PIUS	.39 (.28)	.37 (.22)	.31 (.24)	.39 (.36)				
Excessive Use	.27 (.14)	.25 (.10)	.24 (.11)	.25 (.21)	.78 (.77)			
Social Benefit	.42 (29)	.38 (.23)	.35 (.25)	.40 (.35)	.89 (.85)	.61 (.51)		
Negative Use	.37 (.28)	.34 (.21)	.27 (.24)	.37 (.35)	.96 (.95)	.66 (.63)	.79 (.71)	
Mean	2.59 (2.50)	2.61 (2.50)	2.79 (2.76)	2.31 (2.18)	2.00 (1.67)	2.64 (2.39)	1.99 (1.64)	1.78 (1.44)
Standard deviation	.66 (.66)	.77 (.80)	.67 (.77)	.74 (.77)	.76 (.61)	.88 (.87)	.78 (.65)	.85 (.65)
Alpha	.92 (.94)	.86 (.89)	.76 (.84)	.75 (.80)	.95 (.94)	.71 (.76)	.85 (.83)	.94 (.93)

[a] r≥.15, P=.01; .12< r<.14, P=.05.

[b] within parentheses are values for women.

Results

Descriptive Findings

Descriptive statistics for the SAS and the PIUS for men and women were computed and reported in Table 2. Subscale scores were divided by the number of items in the respective subscales to derive subscale means independent of the number of items. Results show that in social anxiety, both men (mean 2.79, SD 0.67) and women (mean 2.76, SD 0.77) scored the highest on the criticism anxiety subscale and both men (mean 2.31, SD 0.74) and women (mean 2.18, SD 0.77) scored the lowest on the self-depreciation subscale. However, mean differences between men and women were not statistically significant (λ=.02, $F_{3,501}$=2.47, P=.06). Therefore, results failed to support the first hypothesis. On all three PIU dimensions, men scored higher than women (Table 2). MANOVA shows that multivariate difference was statistically significant (λ=.94, $F_{3,501}$=10.69, P<.001). Thus, these results support the second hypothesis.

Canonical Findings

Before canonical correlation analyses, multivariate outliers were screened by Mahalanobis distances. Any Mahalanobis distance greater than 22 (χ^2_6=22.5, P<.001) was considered to be a multivariate outlier. A total of five cases were found to be multivariate outliers and excluded from the further analyses. After multivariate outliers were deleted, Mardia coefficient provided an indication that the data were free from serious violations of multivariate normality (Mardia coefficient=6.57; normalized estimate=7.54). Multivariate homoscedasticity was investigated by Box M statistics, and results show that observed covariance matrices of the SAS and the PIUS were not equal across genders (Box M=45.53, $F_{21,894734}$=2.14, P=.002).

Two separate canonical correlation analyses were performed, one for men and one for women, to investigate multivariate relationships between social anxiety variables (ie, social avoidance, criticism anxiety, and self-depreciation) and PIU variables (ie, excessive use, social benefit, and negative use). Table 2 presents the Pearson product-moment correlation matrix from which canonical roots were generated. Three canonical functions were computed for men, and three canonical functions were computed for women. The strength of the relationship was assessed by the magnitude of the canonical correlation coefficients. Of the three canonical correlation functions computed for men, only the first was significant (R_c=.43, λ=.78, χ^2_9=64.7, P<.001) and accounted for 19% of the overlapping variance. Similarly, only the first canonical function was significant for women (R_c=.36, λ=.87, χ^2_9=33.9, P<.001), which accounted for 13% of the overlapping variance. Therefore, canonical correlation results support the third hypothesis for both men and women. For both groups, standardized canonical coefficients, the percentage of variances accounted in each set, and redundancies were computed and reported in Table 3. The proportion of shared variance together with redundancy indicates that the first canonical variate was moderately related for men. The canonical variate accounted for 64% of the variability in the social anxiety set and 53% in the PIU set. For women, the canonical variate accounted for 41% of the variability in the social anxiety set and 50% in the PIU set.

Table 3. Canonical correlation results.

Variables	First canonical variate			
	Men		Women	
Social anxiety set	r_s [a]	Coefficient[b]	r_s [a]	Coefficient[b]
Social avoidance	−.90	−.39	−.69	.06
Criticism anxiety	−.81	−.11	−.67	.22
Self-depreciation	−.95	−.60	−.99	−1.20
Variance percentage	.80		.64	
Redundancy	.15		.08	
Problematic Internet use set				
Excessive use	−.64	−.02	−.63	.01
Social benefit	−.98	−.70	−.93	−.53
Negative use	−.91	−.36	−.94	−.55
Variance percentage	.73		.71	
Redundancy	.14		.09	
Canonical correlation coefficient (R_c)	.43		.36	
R_c^2	19%		13%	

[a] Structure coefficients (canonical loadings).

[b] Standardized canonical coefficients.

A cut-off score of .30 is suggested as minimum acceptable loading value (ie, structure coefficients) in interpreting the variables of a given canonical variate. All structure coefficients in all sets for both men and women exceeded the acceptable cut-off limits (Table 3). For both men and women, the highest canonical loading in the social anxiety set was self-depreciation (−.95 and −.99, respectively) and the lowest was criticism anxiety (−.81 and −.67, respectively). For men, the highest canonical loading in the PIU set was social benefit (−.98), whereas it was negative use (−.94) for women. For both men and women, the lowest PIU was excessive use (−.64 and −.63, respectively). The pair of canonical variates shows that higher self-depreciation, social avoidance, and criticism anxiety are most associated with greater social benefits for men. However, for women, higher self-depreciation, criticism anxiety, and social avoidance are most associated with greater negative use.

Discussion

Principal Findings

Descriptive results revealed that even though men scored higher than women in the components of social anxiety, multivariate analyses showed that differences were not significant. Therefore, we rejected the first hypothesis. In general, social anxiety studies show gender differences where women suffer more from social anxiety (eg, [43]); however, given the contemporary era, culture, and the characteristics of the present population, such gender differences have reduced to a point where they are not significant anymore. This is one of the major contributions of this study to the existing literature. Our findings support those studies that did not find any significant gender differences in social anxiety (eg, [36]). One plausible explanation for the nonsignificant gender difference in social anxiety is that contemporary gender roles have become more similar in recent years. Enhanced educational opportunities for women or their increasing role in the society have lead women to become more active and thus closed the gap in social anxiety levels between men and women. However, this conclusion needs further support by future studies.

Significant differences were found in PIU. Men’s averages were higher than those of women’s averages in excessive use, social benefit, and negative use. Even though some studies reveal the contrary or no difference, there is stronger research evidence that men tend to experience higher levels of PIU (Table 1). We aimed to clarify such discrepancies in the literature. Our results continue to support the earlier as well as more recent research findings showing higher PIU among men. We conclude that men are under more risk of PIU. Majority of the current studies in the literature have investigated PIU as a unidimensional construct (eg, [13,48]). However, one of the unique contributions of this research is that it viewed PIU as a multidimensional construct. We found that men showed more difficulties than women in terms of running away from personal problems (ie, social benefit), used the Internet more excessively, and experienced interpersonal problems with significant others due to Internet use. Therefore, we conclude that men are under greater risk of social impairments due to PIU. These results support gender schema theory, which assert that gender affects individuals’ cognitions and behaviors not only in conventional settings but also on the Internet. We can say that our results extend the implications of gender schema theory and social role theory to the virtual domain.

This study is the first in the literature that considered the multivariate nature of both social anxiety and PIU in studying the relationship between the two. We found significant multivariate relationships between sets social anxiety variables and sets of PIU variables for both men and women and thus supported the third hypothesis. These relationships support the cognitive-behavioral model of PIU and indicate that psychological health (ie, social anxiety) and certain problematic behaviors on the Internet are associated to each other. Also supported by the findings is the displacement hypothesis. As the levels of social anxiety increase and people spend less time in real-life relationships and more time on the Internet, risk of developing PIU becomes greater.

This study found specific relationships between certain aspects of social anxiety and certain aspects of PIU for men and women. For example, social anxiety components are mostly related to social benefit problems among men; however, they are most relevant to negative use problems among women. These results support the finding of Hetzel-Riggin and Pritchard [44], which show that men use the Internet more for seeking social support. Finally, results also lend support to the lonely drawn to the Internet hypothesis, indicating that greater social problems may lead to higher Internet use.

The strength of the association between the two constructs can be assessed by squared multiple correlations, which show the amount of shared variability. In this study, 19% and 13% variability were shared between social anxiety and PIU for men and women, respectively. According to guidelines suggested by Cohen [54], these effect sizes can be considered large. The magnitudes of highly significant canonical correlations (averaged R of .40) indicate that there is a substantial amount of association between social anxiety and PIU. It can also be concluded that association between social anxiety and PIU is stronger for men than it is for women.

Results indicate that men or women who are more self-depreciative, who avoid social contact, and who suffer from higher levels of criticism anxiety tend to use the Internet more for social benefit and engage in more negative and excessive use. Similar to the findings of this study, Peter et al [55] reported that Internet communication is highly valued by socially anxious adolescents. In sum, we conclude that social anxiety is an important factor in PIU. People with higher levels of social anxiety may fulfill their needs on the Internet and seem to prefer Web-based communication, but the patterns of such a fulfillment vary between men and women.

We also found that both social anxiety and PIU canonical variates were strongly related to all the variables that form their respected variates. The proportions of variances and redundancies indicate that the social anxiety canonical variate was stronger in men than in women. In PIU, canonical variates were almost identical among men and women. For both genders, self-depreciation was the highest loading variable in social anxiety. This finding can be interpreted to mean that being unhappy with oneself contributes the most to social anxiety. However, in PIU, social benefit contributes the most to social anxiety for men, but for women, negative use contributes the most to social anxiety. Additionally, these significant relationships were invariant across men and women even though the association was stronger for men than it was for women.

Implications

A number of implications from the findings of the study are to be mentioned at this point. When designing intervention programs in dealing with PIU, it should be kept in mind that interventions need to use different approaches based on the varying needs of men and women identified in this research. In addition, future research should continue to track whether gender differences in social anxiety continue to shrink. We advise that future research continue to investigate PIU and social anxiety as multidimensional constructs.

Limitations

Finally, this study has a few limitations. First, both social anxiety and PIU are shown to be related to many different variables in the literature, most of which are not included in this study. This study is limited to specific relationships and gender differences. Second, the study was conducted with a nonclinical sample of college students. Therefore, results are only generalizable to similar populations. Third, this study relied on self-report measures and thus might suffer from common method bias. Other means of assessing social anxiety and PIU might reveal different results. Fourth, results show associations between the constructs under investigation, but they do not imply any causality. Thus, it would be inappropriate to conclude that social anxiety *causes* PIU or vice versa.

Conflicts of Interest

None declared.

References

1. Huang C. Internet use and psychological well-being: a meta-analysis. Cyberpsychol Behav Soc Net 2010 Jun;13(3):241-249. [doi: 10.1089/cyber.2009.0217]
2. Niemz K, Griffiths M, Banyard P. Prevalence of pathological Internet use among university students and correlations with self-esteem, the General Health Questionnaire (GHQ), and disinhibition. Cyberpsychol Behav 2005 Dec;8(6):562-570. [doi: 10.1089/cpb.2005.8.562] [Medline: 16332167]
3. Suhail K, Bargees Z. Effects of excessive Internet use on undergraduate students in Pakistan. Cyberpsychol Behav 2006 Jun;9(3):297-307. [doi: 10.1089/cpb.2006.9.297] [Medline: 16780397]
4. Davis RA. A cognitive-behavioral model of pathological Internet use. Comput Human Behav 2001;17(2):187-195. [doi: 10.1016/S0747-5632(00)00041-8]
5. Beard KW, Wolf EM. Modification in the proposed diagnostic criteria for Internet addiction. Cyberpsychol Behav 2001 Jun;4(3):377-383. [doi: 10.1089/109493101300210286] [Medline: 11710263]
6. Ostovar S, Allahyar N, Aminpoor H, Moafian F, Nor Md MB, Griffiths MD. Internet addiction and its psychosocial risks (depression, anxiety, stress and loneliness) among Iranian adolescents and young adults: a structural equation model in a cross-sectional study. Int J Ment Health Addiction 2016;14(3):257-267. [doi: 10.1007/s11469-015-9628-0]
7. Yu S, Wu AMS, Pesigan IJA. Cognitive and psychological health risk factors of social networking addiction. Int J Ment Health Addiction 2016;14(4):550-564. [doi: 10.1007/s11469-015-9612-8]
8. Koronczai B, Kökönyei G, Urbán R, Király O, Nagygyörgy K, Felvinczi K, et al. Confirmation of the Chinese version of the Problematic Internet Use Questionnaire Short Form (PIUQ-SF). Int J Ment Health Addiction 2017 Feb;15(1):191-197. [doi: 10.1007/s11469-016-9664-4]
9. Bischof-Kastner C, Kuntsche E, Wolstein J. Identifying problematic Internet users: development and validation of the Internet Motive Questionnaire for Adolescents (IMQ-A). J Med Int Res 2014 Oct;16(10):e230. [Medline: 4210952]
10. Faraci P, Craparo G, Messina R, Severino S. Internet Addiction Test (IAT): which is the best factorial solution? J Med Internet Res 2013 Oct 9;15(10):e225 [FREE Full text] [Medline: 24184961]
11. Wellman B, Haythornthwite C, editors. Internet use, interpersonal relations, and sociability. In: The Internet in Everyday Life. Malden, MA: Blackwell; 2002:215-243.
12. Hu M. Will online chat help alleviate mood loneliness? Cyberpsychol Behav 2009 Apr;12(2):219-223. [doi: 10.1089/cpb.2008.0134] [Medline: 19250012]
13. Caplan SE. Relations among loneliness, social anxiety, and problematic Internet use. Cyberpsychol Behav 2007 Apr;10(2):234-242. [doi: 10.1089/cpb.2006.9963] [Medline: 17474841]
14. Bonetti L, Campbell MA, Gilmore L. The relationship of loneliness and social anxiety with children's and adolescents' online communication. Cyberpsychol Behav Soc Net 2010 Jun;13(3):279-285. [doi: 10.1089/cyber.2009.0215]
15. Eichenberg C, Schott M, Decker O, Sindelar B. Attachment style and Internet addiction: an online survey. J Med Int Res 2017 May 17;19(5):e170. [Medline: 28526662]
16. APA. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Washington, DC: American Psychiatric Association Publishing; 2013.

17. National Institute of Mental Health. Social phobia (social anxiety disorder): always embarrassed. Bethesda: National Institute of Mental Health; 2010.
18. Weinstein A, Dorani D, Elhadif R, Bukovza Y, Yarmulnik A, Dannon P. Internet addiction is associated with social anxiety in young adults. Ann Clin Psychiatry 2015 Feb;27(1):4-9. [Medline: 25696775]
19. Fehm L, Pelissolo A, Furmark T, Wittchen HU. Size and burden of social phobia in Europe. Eur Neuropsychopharmacol 2005 Aug;15(4):453-462. [doi: 10.1016/j.euroneuro.2005.04.002] [Medline: 15921898]
20. Chiang HS, Hsiao KL. YouTube stickiness: the needs, personal, and environmental perspective. Internet Research 2015 Feb;25(1):85-106. [doi: 10.1108/IntR-11-2013-0236]
21. Sheeks MS, Birchmeier ZP. Shyness, sociability, and the use of computer-mediated communication in relationship development. Cyberpsychol Behav 2007 Feb;10(1):64-70. [doi: 10.1089/cpb.2006.9991] [Medline: 17305450]
22. Chow PI, Fua K, Huang Y, Bonelli W, Xiong H, Barnes LE, et al. Using mobile sensing to test clinical models of depression, social anxiety, state affect, and social isolation among college students. J Med Int Res 2017 Mar;19(3):e62. [Medline: 5357317]
23. Hedman E, Furmark T, Carlbring P, Ljótsson B, Rück C, Lindefors N, et al. A 5-Year follow-up of internet-based cognitive behavior therapy for social anxiety disorder. J Med Internet Res 2011 Jun 15;13(2):e39 [FREE Full text] [doi: 10.2196/jmir.1776] [Medline: 21676694]
24. Bussey K, Bandura A. Social cognitive theory of gender development and differentiation. Psychol Rev 1999;106(4):676-713. [doi: 10.1037/0033-295X.106.4.676]
25. Eagly AH. Sex Differences in Social Behavior: A Social Role Interpretation. New Jersey: Lawrence Erlbaum Associates, Inc; 1987.
26. Bem SL. The measurement of psychological androgyny. J Consult Clin Psyc 1974;42(2):155-162. [doi: 10.1037/h0036215]
27. Bem SL, Martyna W, Watson C. Sex typing and androgyny: further explorations of the expressive domain. J Pers Soc Psychol 1976 Nov;34(5):1016-1023. [Medline: 993980]
28. Bem SL, Lewis SA. Sex role adaptability: one consequence of psychological androgyny. J Person Soc Psyc 1975;31(4):634-643. [doi: 10.1037/h0077098]
29. Bem SL. On the utility of alternative procedures for assessing psychological androgyny. J Consult Clin Psyc 1977;45(2):196-205. [doi: 10.1037/0022-006X.45.2.196]
30. Bem SL. Gender schema theory: a cognitive account of sex typing. Psychol Rev 1981;88(4):354-364. [doi: 10.1037/0033-295X.88.4.354]
31. Bem SL. Gender schema theory and its implications for child development: raising gender-aschematic children in a gender-schematic society. Signs 1983;8(4):598-616. [doi: 10.1086/493998]
32. Uzzell D, Horne N. The influence of biological sex, sexuality and gender role on interpersonal distance. Brit J Soc Psyc 2006;45(Pt 3):579-597. [Medline: 16984722]
33. Eagly AH, Wood W, Diekman AH. Social role theory of sex differences and similarities: a current appraisal. In: Eckes T, Trautner HM, editors. The Developmental Social Psychology of Gender. Mahwah, NJ: Erlbaum; 2000:123-174.
34. Eagly AH, Wood W. Social role theory. In: Van Lange PA, Kruglanski A, Higgins ET, editors. Handbook of Theories in Social Psychology (Vol. 2). London, England: Sage; 2011:458-476.
35. McLean CP, Asnaani A, Litz BT, Hofmann SG. Gender differences in anxiety disorders: prevalence, course of illness, comorbidity and burden of illness. J Psychiatr Res 2011 Aug;45(8):1027-1035 [FREE Full text] [doi: 10.1016/j.jpsychires.2011.03.006] [Medline: 21439576]
36. Bujala A. Gender differences in Internet usage. Act Univers Lodzen Folia Sociolo 2012;43:49-67. [doi: 10.18778/0208-600X]
37. Spielberger CD. Anxiety: Current Trends in Theory and Research. New York: Academic Press; 1972:2-23.
38. Baloğlu M. Individual differences in statistics anxiety among college students. Per Ind Dif 2003 Apr;34(5):855-865. [doi: 10.1016/S0191-8869(02)00076-4]
39. Baloğlu M, Koçak R. A multivariate investigation of the differences in mathematics anxiety. Per Ind Dif 2006 May;40(7):1325-1335. [doi: 10.1016/j.paid.2005.10.009]
40. Sarıkaya Y, Baloğlu M. The development and psychometric properties of the Turkish Death Anxiety Scale (TDAS). Death Stud 2016 Aug;40(7):419-431. [Medline: 26933746]
41. Erath SA, Flanagan KS, Bierman KL. Social anxiety and peer relations in early adolescence: behavioral and cognitive factors. J Abnorm Child Psychol 2007 Jun;35(3):405-416. [doi: 10.1007/s10802-007-9099-2] [Medline: 17265192]
42. Rana SA, Akhtar S, Tahir MA. Parenting styles and social anxiety among adolescents. New Horiz 2013;7(2).
43. Cao H, Sun Y, Wan Y, Hao J, Tao F. Problematic Internet use in Chinese adolescents and its relation to psychosomatic symptoms and life satisfaction. BMC Public Health 2011 Oct 14;11:802 [FREE Full text] [doi: 10.1186/1471-2458-11-802] [Medline: 21995654]
44. Hetzel-Riggin MD, Pritchard JR. Predicting problematic Internet use in men and women: the contributions of psychological distress, coping style, and body esteem. Cyberpsychol Behav Soc Netw 2011 Sep;14(9):519-525. [doi: 10.1089/cyber.2010.0314] [Medline: 21342011]

45. Kormas G, Critselis E, Janikian M, Kafetzis D, Tsitsika A. Risk factors and psychosocial characteristics of potential problematic and problematic Internet use among adolescents: a cross-sectional study. BMC Public Health 2011 Jul 27;11:595 [FREE Full text] [doi: 10.1186/1471-2458-11-595] [Medline: 21794167]
46. Mottram AJ, Fleming MJ. Extraversion, impulsivity, and online group membership as predictors of problematic Internet use. Cyberpsychol Behav 2009 Jun;12(3):319-321. [doi: 10.1089/cpb.2007.0170] [Medline: 19445635]
47. Durkee T, Kaess M, Carli V, Parzer P, Wasserman C, Floderus B, et al. Prevalence of pathological Internet use among adolescents in Europe: demographic and social factors. Addiction 2012 Dec;107(12):2210-2222. [doi: 10.1111/j.1360-0443.2012.03946.x] [Medline: 22621402]
48. Schimmenti A, Passanisi A, Gervasi AM, Manzella S, Famà FI. Insecure attachment attitudes in the onset of problematic Internet use among late adolescents. Child Psychiatry Hum Dev 2014 Oct;45(5):588-595. [doi: 10.1007/s10578-013-0428-0] [Medline: 24338269]
49. La Greca AM, Lopez N. Social anxiety among adolescents: linkages with peer relations friendships. J Abnor Child Psych 1998;26(2):83-94. [doi: 10.1023/A:1022684520514]
50. Weiser EB. Gender differences in Internet use patterns and Internet application preferences: a two-sample comparison. Cyberpsychol Behav 2004;3(2):167-178. [doi: 10.1089/109493100316012]
51. Lin MP, Ko HC, Wu JY. Prevalence and psychosocial risk factors associated with Internet addiction in a nationally representative sample of college students in Taiwan. Cyberpsychol Behav Soc Netw 2011 Dec;14(12):741-746. [doi: 10.1089/cyber.2010.0574]
52. Özbay Y, Palancı M. The validity and reliability of the Social Anxiety Scale for university students. 2001 Sep 10 Presented at: 6th Nat Cong Psyc Coun Guidance; September 2001; Ankara, Turkey.
53. Ceyhan E, Ceyhan AA, Gürcan A. The validity and reliability of the Problematic Internet Usage Scale. Educ Sci Theory Pract 2007;7(1):411-416.
54. Cohen J. Statistical Power Analysis for the Behavioral Sciences. New Jersey: Hillsdale; 1988.
55. Peter J, Valkenburg PM, Schouten AP. Developing a model of adolescent friendship formation on the Internet. Cyberpsychol Behav 2005;8(5):423-430. [Medline: 16232035]

Abbreviations

DSM: Diagnostic and Statistical Manual of Mental Disorders
MANOVA: multivariate analysis of variance
PIU: problematic Internet use
PIUS: Problematic Internet Use Scale
SAS: Social Anxiety Scale

Edited by J Torous; submitted 12.09.17; peer-reviewed by M Ak, V Santos; comments to author 13.10.17; accepted 20.10.17; published 24.01.18.

Please cite as:
Baloğlu M, Özteke Kozan Hİ, Kesici Ş
Gender Differences in and the Relationships Between Social Anxiety and Problematic Internet Use: Canonical Analysis
J Med Internet Res 2018;20(1):e33
URL: http://www.jmir.org/2018/1/e33/
doi:10.2196/jmir.8947
PMID:29367182

Original Paper

Evidence From Web-Based Dietary Search Patterns to the Role of B12 Deficiency in Non-Specific Chronic Pain: A Large-Scale Observational Study

Eitan Giat[1], MD, PhD; Elad Yom-Tov[2], PhD

[1]Rheumatology Unit, The Autoimmune Center, Sheba Medical Center, Ramat Gan, Israel
[2]Microsoft Research, Herzeliya, Israel

Corresponding Author:
Elad Yom-Tov, PhD
Microsoft Research
13 Shenkar St
Herzeliya, 46875
Israel
Phone: 972 747111359
Email: eladyt@yahoo.com

Abstract

Background: Profound vitamin B12 deficiency is a known cause of disease, but the role of low or intermediate levels of B12 in the development of neuropathy and other neuropsychiatric symptoms, as well as the relationship between eating meat and B12 levels, is unclear.

Objective: The objective of our study was to investigate the role of low or intermediate levels of B12 in the development of neuropathy and other neuropsychiatric symptoms.

Methods: We used food-related Internet search patterns from a sample of 8.5 million people based in the US as a proxy for B12 intake and correlated these searches with Internet searches related to possible effects of B12 deficiency.

Results: Food-related search patterns were highly correlated with known consumption and food-related searches (ρ=.69). Awareness of B12 deficiency was associated with a higher consumption of B12-rich foods and with queries for B12 supplements. Searches for terms related to neurological disorders were correlated with searches for B12-poor foods, in contrast with control terms. Popular medicines, those having fewer indications, and those which are predominantly used to treat pain, were more strongly correlated with the ability to predict neuropathic pain queries using the B12 contents of food.

Conclusions: Our findings show that Internet search patterns are a useful way of investigating health questions in large populations, and suggest that low B12 intake may be associated with a broader spectrum of neurological disorders than previously thought.

(J Med Internet Res 2018;20(1):e4) doi:10.2196/jmir.8667

KEYWORDS

B12 deficiency; diet; Internet searches; neuropsychiatric symptoms; neuropathy

Introduction

Vitamin B12 has long been known to be a cause of megaloblastic anemia and neurological disorders [1]. Very low levels of B12 cause subacute combined degeneration of the dorsal and lateral spinal columns, manifested by paresthesia, ataxia, impaired sense of vibration and proprioception, progressive weakness, spasticity, and paraplegia. These symptoms may be irreversible, depending on duration. Other neurological abnormalities may be caused by B12 deficiency, such as peripheral neuropathy, memory loss, irritability, and dementia [2,3], but the association of these symptoms with serum B12 levels and B12 intake is unclear. Many cases of B12 deficiency come to medical attention because of these symptoms, but seemingly asymptomatic B12 deficiency is very common, with prevalences up to 10% to 25% [4]. Furthermore, neurological symptoms due to B12 deficiency may appear in patients within a normal range of serum B12 levels [5].

There is no clear cut-off for serum B12 levels [6]—levels higher than 300 pg/ml are considered normal, whereas between 200 and 300 pg/ml are considered borderline, and less than 200 pg/ml are considered low. In addition, measurements of serum B12 levels are not always reliable [6]. Several methods are used by different labs, resulting in different normal ranges. Furthermore, results are highly variable, with absolute intra-individual variation less than 100 pg/ml on repeat testing in one fifth of the patients. Serum B12 concentration may be normal in 5% of patients with documented B12 deficiency [7].

Intermediate or low levels of B12 may contribute to the development of symptoms that are not typically associated with B12 deficiency. For instance, B12 dietary intake inversely correlated with sleep duration in healthy Japanese patients [8]. B12 levels also inversely correlated with cardiovascular autonomic neuropathy in diabetics (mean level of 289 pg/ml), suggesting that intermediate levels may also be clinically significant. Some have suggested low B12 levels may be a risk factor for depression, but evidence is conflicting [9,10].

The effect of diet on B12 levels is not fully understood and depends on food source and bioavailability. Usual western diets contain between 5 to 7 mcg of daily B12 intake, which is similar to the daily recommended intake, while body stores of B12 are 2000 to 5000 mcg [6]. While foods derived from animals are the major dietary sources of B12, supplements containing B12 are commonly used in the United States and many foods are fortified with B12 (eg, non-dairy milks, meat substitutes, and breakfast cereals) [4]. B12 is naturally protein-bound and its absorption depends on gastric acidity and intrinsic factors, which may be influenced by drugs and comorbidities; reduced absorption of unbound B12 found in vitamin supplements is common among healthy patients. Evidence on the effect of B12 intake from meat is confounding. A study assessing 2999 participants from the Framingham Offspring study found meat intake did not affect B12 levels and suggested a protective effect from supplements and B12 fortified foods [4]. In Norway, where cereals are not fortified with B12, dietary intake of B12 from milk and fish was associated with higher plasma vitamin B12 concentrations, whereas B12 from meat or eggs intake was not [11]. In contrary, analysis from the Dutch B-PROOF study [12] suggested that the bioavailability of B12 intake from meat is comparable to B12 intake from milk. To our knowledge, B12 intakes from different types of meat, such as beef, pork or chicken have not been compared.

Internet search activity has been shown to reflect behavior in the physical world [13] since people use such media for information regarding many everyday behaviors, reflecting their diet, habits, and illnesses [14]. In this study, we used these data to assess food consumption and physical symptoms experienced by people using search engines. With respect to food consumption, past work has found a high correlation between recipe searches and actual consumption [15]. The caloric values of recipes saved (bookmarked) on a website [16] and on Twitter [17] were found to be correlated with obesity at US state and county levels. The link between medical symptoms experienced by people and their searches has been used to understand a range of medical phenomena, including the adverse reactions of medical drugs [18], early detection of cancer [19], and precursors to disease [20].

Many symptoms typically associated with B12 deficiency can occur among patients with intermediate to normal levels of B12. We therefore sought to assess the impact of diet on these symptoms. Here, we analyzed Internet searches related to food to quantify dietary B12 intake from different sources of food and their association with B12-related symptoms. We also sought to find a connection between B12 intake and other symptoms that are not typically associated with B12 deficiency, such as chronic pain or fibromyalgia.

Methods

A list of 212 target terms related to different types of chronic pain were identified from the following categories: (1) antidepressants, (2) neuropathic drugs, (3) other pharmaceuticals (antihistamines and anti-acids), (4) descriptions of pain, (5) disorders related to excess acid, (6) over the counter antacids, (7) psychotherapy, and (8) medical cannabis. The terms are listed in Multimedia Appendix 1. Generic and commercial drug names of these disorders were taken from the Uptodate, Micromedex and drugs.com websites. We also tested 27 medical terms (pharmaceuticals and conditions) which were unlikely to be associated with B12 deficiency and used these as control terms. Control terms were chosen from an online medical dictionary (Medterms). To ensure randomness, we chose the first terms in the alphabetical list, omitting terms which were rarely searched and could bias the results.

We extracted all searches made in English by people in the United States on the Bing search engine during October 2016. For each search, we extracted an anonymized user identifier, the text of the search, and the zip code from which the user made the search. The classes of search categories were (1) recipe searches, searches that contained a reference to a food recipe, as identified by a propriety classifier; (2) term searches, searches that contained 1 or more of the target terms (212 terms above) or the 27 control terms; and (3) B12 searches, searches that contained the term “B12” in them.

To identify ingredients required for each recipe we used the list of recipes from Recipes Wikia [21]. Specifically, we extracted all recipes that contained 1 or more of the following 12 food ingredients (food types), which have a variety of B12 contents: shellfish, mackerel, trout, salmon, tuna, pig, beef, turkey, chicken, egg, milk, and tomatoes. A total of 9449 recipes contained 1 or more of these ingredients.

Additional data required for analysis included the following: (1) B12 contents per serving of the 12 food ingredients, from the US National Institutes of Health Office of Dietary Supplements (Table 1); (2) food spending data from the US Bureau of Labor Statistic’s Consumer Expenditure Survey 2015; (3) cost of food items from the US Department of Agriculture’s Center for Nutrition Policy and Promotion; (4) food consumption data from the US Department of Agriculture’s Economic Research Service; and (5) the number of indications for each medical drug from the Canadian Vigilance Adverse Reactions Online database. All indications reported by at least 10 cases were considered.

Table 1. B12 content in different types of meat according to the National Institutes of Health Office of Dietary Supplements.

Meat source	B12 content, mcg/100 grams
Shellfish	98.9
Mackerel	19
Beef	6
Trout	3.8
Salmon	2.4
Tuna	1.6
Milk	0.9
Turkey	0.8
Egg	0.6
Pork	0.4
Chicken	0.3

The study was deemed exempt by the Microsoft Institutional Review Board.

A demographic analysis of people searching for B12 is provided in Multimedia Appendix 2. The distribution of food searches by US state is given in Multimedia Appendix 3.

Results

Approximately 8.4 million people searched for a target medical term; the term “B12” or a food recipe was included in the analysis. Of the 212 target medical terms, only 101 were queried by 1000 or more people and were included in the analysis. A total of 3297 recipes were matched to at least 1 query. Manual examination of a random sample of 400 queries found that 87.8% (351/400) of queries were correctly matched to a recipe.

Here, we used searches for food recipes as a proxy for consumption, as suggested by West et al [15]. To further validate that this proxy is accurate, we correlated the fraction of known expenditure on 6 food items (beef, pork, poultry, fish and seafood, eggs, and milk) at each of 4 US regions (Northeast, Midwest, South, and West) reported by the US Bureau of Labor Statistic’s Consumer Expenditure Survey 2015 (see Methods), with the number of queries for recipes containing each of these food items weighted by their cost. The Pearson correlation between the 2 was 0.69 (*P*=0.0002, N=24), suggesting that searches for foods are strongly correlated with consumption. Thus, we posited that B12 consumption can be estimated through recipe searches.

Awareness of B12 deficiency is also correlated with consumption. The estimated B12 consumption of users was computed by summing the multiplication of the B12 contents per serving of each of the 12 food types by the number of times a user searched about each food type. The average value of the estimated consumption for people who asked about B12 (people who made B12 searches) was 2.407 mcg, compared to 2.395 mcg for those who did not, resulting in a 0.5% difference and a statistically significant ranksum test value (*P*<.0001). However, the difference in estimated consumption for the 4701 people who specifically asked about B12 deficiency was 2.219 mcg compared to 2.395 mcg for those who did not. This correlated to an 8% difference and a ranksum test value of *P*<.0001. We interpreted this finding as additional supporting evidence for the claim that estimating B12 consumption through searches for recipes is valid, given the lower estimated B12 consumption for people who indicated an awareness of this deficiency.

The vast majority (99.94%, 8,429,637/8,434,338) of people in our sample did not ask about B12 deficiency, supplements, or serum level. However, among people who asked about B12 deficiency, 3.83% (180/4701) asked about B12 supplements (either in the form of dietary supplements or injections), compared to 0.03% (2132/8,429,637) in the population who didn’t ask about B12 deficiency. Thus, approximately 151 more times people who asked about deficiency also asked about supplements.

To estimate the correlation between food consumption and the likelihood of asking about each of the medical terms, we constructed an individual-level model where each person is represented through the number of searches they made for recipes containing each of 12 food items and, separately, as to whether they asked about each of the target terms. The behaviors were linked through a linear classification model where the independent terms were the number of searches for each food item and the dependent variables were whether a person asked about the medical term.

The 10 medical terms for which the individual-level model reached the highest goodness of fit (R^2) value (denoted by R^2I), together with the correlation between model coefficients (per food item) and the B12 contents per serving of each food item, denoted by CoB12, are shown in Table 2. The regression coefficients for computing CoB12 are shown in Multimedia Appendix 4. R^2I values were low even for the best fitting models, but their median value was 78 times greater than R^2I values for the control terms (medical terms unassociated with B12 deficiency for the top 10 terms in each list; ranksum *P*=0.002). We further noted that R^2I was correlated with the number of people who asked about each medical condition (ρ=.86, *P*<.0001). Thus, conditions we hypothesized may be linked to B12 deficiency can indeed be associated with food consumption, especially for the more common medical conditions.

Next, we used the models constructed above to estimate the contribution of B12 contents to each medical condition. The correlation between model coefficients and B12 contents of food items (CoB12) was negative for all terms in Table 2, indicating that people searching for B12-rich foods were less likely to search for the medical terms, compared to people searching for B12-poor foods.

Different CoB12 values observed across medical terms can be explained by several factors. First, the average values of the regression coefficients for all but tomatoes were at least 2.3 times greater than those of tomatoes, whose B12 content is negligible. Second, the Spearman correlation between R^2I and CoB12 was –.29 (*P*=0.004, N=101), whereas for the control terms it was non-significant (ρ=.35, *P*=0.08, N=27). Thus, among the target terms, coefficients of better fitting models were larger (more negative) than of poorly fitting models. Finally, CoB12 for medical drug terms was correlated with drug indications, as described below. On average, the correlation for the control terms was only .04, whereas the same number for the target medical terms was, on average –.31 (ranksum *P*=.01). Moreover, positive correlations were only observed for individual control terms and never among the target terms.

The number of possible indications of a drug can affect its correlation with CoB12, since indications to unrelated conditions may mask the true correlation. We identified the number of indications from the Canadian Adverse Reactions database. Indications were further stratified to identify those indications that contained the term “pain.” The number of indications was strongly correlated with the number of people who asked about a medical condition (ρ=.53, *P*<.0001, N=95), but the number of pain indications only marginally significantly correlated with the number of people (ρ=.29, *P*=0.048, N=95). These correlations were likely due to broad-spectrum drugs being offered to treat more conditions than drugs which can only be prescribed to specific indications.

Next, we tested the correlation between the broadness of an indication of a drug on CoB12. We modeled CoB12 of medical drugs as a function of the interaction between the number of people who asked about a term, the number of indications, and the number of pain-related indications using a rank regression model, weighting samples by R^2I. The model reached an R^2 of 0.48 (*P*<.0001, N=49). Statistically significantly correlated coefficients of the model are shown in Textbox 2.

Table 2. R^2 of the individual-level models to predict the likelihood of asking about medical terms given questions about foods (denoted by R^2I) and the correlation between model coefficients and B12 contents of these food items for the 10 medical terms with the highest R^2I values.

Medical term	R^2I[a]	CoB12[b]
Gabapentin	0.003	–0.405
Tramadol	0.003	–0.473
Neuropathy	0.002	–0.322
Omeprazole	0.002	–0.397
Sertraline	0.002	–0.303
Citalopram	0.002	–0.333
Oxycodone	0.002	–0.269
Duloxetine	0.002	–0.329
Trazodone	0.002	–0.354

[a] R^2I: R^2 of the individual-level models to predict the likelihood of asking about medical terms given questions about foods.

[b] CoB12: correlation between model coefficients (per food item) and the B12 contents per serving of each food item.

Textbox 2. Statistically significantly correlated coefficients of the model.

The number of indications:

- Slope was –0.43 (*P*<.0001)
- Meaning that fewer indications were correlated with higher CoB12

The number of pain indications:

- Slope was 0.20 (*P*=0.04)
- Meaning that the more pain-related indications, the higher CoB12

We interpreted the first as showing that the CoB12 of non-specific medicines was lower, as the effect of B12 deficiency was masked by people with other medical conditions. The second correlation was interpreted as showing that the more a drug is used to treat pain (ostensibly caused partly by B12 deficiency), the better the correlation to B12 consumption.

Taken together, our results indicated that the more focused a term is in its medical application, and the more popular, the better the model was in estimating the correlation of B12 consumption with its appearance. The more people consume B12-rich foods, the less likely they were to query for those medical terms we hypothesized were related to B12 deficiency.

Discussion

Principal Findings

Our study is the first to employ an analysis of Internet searches to estimate B12 intake and its effect. We showed a strong correlation between food consumption and Internet searches, indicating that searching for recipes reflected food consumption. Our analysis allowed for an assessment of B12 intake from different types of foods, especially meats. The Framingham Offspring study [4] found that B12 intake from meat did not change B12 levels, but this study regarded all types of meat as one entity. Our data, strongly suggested the different types of meat provided different B12 intakes and therefore should be analyzed separately. Most physicians are aware of the risk of B12 deficiency in strict vegans, but the intake from an average non-vegetarian diet is considered sufficient in B12, with no preference to any type of meat. This is likely because an average non-vegetarian diet includes the recommended daily intake. In addition, vitamin B12 stores are high and easy to replenish. Other causes limiting B12 absorption, such as gastric acidity and intrinsic factor deficiency, are considered much more important factors in B12 deficiency. As a result, physicians do not consider non-vegetarians to be at risk for B12 deficiency, with no discrimination between beef- or poultry-based diets. Our study is the first to assess the differences in B12 intake from different sources of meat in daily diet and suggests that the distinction between vegetarians and non-vegetarians may be inadequate for evaluating the risk for B12 deficiency.

Our data also showed that some people are aware of B12 deficiency and that this was reflected in their diet. However, not all patients with B12 deficiency are aware of this shortage. Indeed, difficulties in assessing B12 levels may cause both the patient and clinician to be unaware of its shortage. Here, we showed that low B12 intake was correlated with symptoms associated with B12 deficiency among people who did not search for B12-rich foods, supposedly unaware of B12 deficiency. We performed most of our analyses on a population of people who did not express awareness of B12 deficiency or sufficient intake thereof because patients who are aware of B12 deficiency tend to use supplemental B12, which may mask the effect of dietary B12 intake. Besides further emphasizing the importance of dietary trends, these results suggest that a significant population of patients may suffer from symptomatic B12 deficiency without being aware of the cause of their symptoms. To date, studies evaluating the effect of B12 intake only measured B12 serum levels as an outcome. Our results suggest that other outcomes, such as peripheral neuropathy and other neurological complaints should also be considered.

Our study found a correlation between B12 intake and different drugs and diseases. Paresthesia, which is neuropathic in origin, may be a symptom of B12 deficiency. Not surprisingly, the 4 strongest correlations were related to neuropathy and pain and its treatment. Gabapentin and Lyrica are common treatments for neuropathic pain or central neuropathic pain. Tramadol is also prescribed for neuropathic pain as well as other types of pain. Internet searches for sertraline, citalopram, duloxetine, and trazodone were also inversely correlated with the estimated B12 intake. These drugs are common anti-depressants and our results strengthen the yet controversial association between low B12 levels and depression [22]. Searches for oxycodone, a potent opiate painkiller, were also associated with lower B12 dietary intake. Oxycodone is not the drug of choice for neuropathic pain, but is usually reserved for severe or refractory pain. B12 is not typically associated with non-neuropathic pain, though parenteral B12 has been shown to alleviate low back pain [23]. The association of B12 intake and oxycodone suggested that B12 may have a role in severe or refractory pain which is not neuropathic. Interestingly, omeprazole, an important proton pump inhibitor, was also correlated with decreased B12 dietary intake. The reason for this relationship is unclear. Omeprazole, through its gastric acid lowering effects, is considered to be a cause for B12 deficiency, but B12 deficiency is not known to cause gastric acidity. A possible explanation might be that peptic pain may be amplified in B12 deficiency, but this has never been shown or looked into.

Limitations

Our study has several limitations. First, a recipe search does not necessarily translate to precise personal consumption, which may have biased the results. We validated our results at the population level (regional consumption), but individual consumption may differ. Past work has found a high correlation between recipe searches and actual consumption [15]. Moreover, in our study we required that recipe searches correlated with the profile of food consumption (ie, the fraction of beef over chicken), not the actual rates of individual consumption. In addition, though correlations were not very high, the fact that despite this limitation we were able to show differences in the use of the target terms as a function of B12 content (but not in searches unrelated to B12) strengthened the notion that differences in B12 intake from different types of meat are important in determining the risk for B12 deficiency and related symptoms.

The use of Internet searches as a proxy for health conditions may also be biased. For example, B12 has been suggested to be associated with depression, but the depression may cause a decrease in Internet use [24], making it difficult to identify this disorder through Internet queries.

Another limitation is our inability to quantify the impact of dietary B12 on serum levels. Our model identified an increase of risk with consumption of meat with low B12 content, but it cannot predict the serum B12 levels for a specific diet. On the other hand, the difficulties in relying on B12 levels and the occurrence of clinical B12 deficiency alongside seemingly normal serum B12 levels, suggest that assessing symptoms may be more important than estimating serum levels.

Our work focused on people in the United States. A comparison to other countries and other dietary preferences could enhance our understanding of the effect of B12 intake on pain.

Despite these limitations, utilizing Internet searches as a proxy for B12 intake has the advantage of overcoming memory bias and has the ability to go into dietary details, which is much more difficult to perform in questionnaire-based studies.

Conclusion

Our data suggest that meat alone is not sufficient to prevent B12 deficiency and that the source of meat should also be considered. Our data also suggests that B12 intake inversely correlates with neurological symptoms, implying a role for B12 among a seemingly unaware population. Physicians should be aware of the possible role of B12 in any patient with neurological complaints or unexplained pain. Further research is necessary to confirm and determine the clinical significance of our results.

Conflicts of Interest

EYT is an employee of Microsoft, owner of the Bing search engine.

Multimedia Appendix 1

Terms used in the study.

[PDF File (Adobe PDF File), 331KB - jmir_v20i1e4_app1.pdf]

Multimedia Appendix 2

Demographics.

[PDF File (Adobe PDF File), 669KB - jmir_v20i1e4_app2.pdf]

Multimedia Appendix 3

Maps of food consumption.

[PDF File (Adobe PDF File), 342KB - jmir_v20i1e4_app3.pdf]

Multimedia Appendix 4

Table 2 with regression coefficients for statistically significant interactions.

[PDF File (Adobe PDF File), 482KB - jmir_v20i1e4_app4.pdf]

References

1. Graham SM, Arvela OM, Wise GA. Long-term neurologic consequences of nutritional vitamin B12 deficiency in infants. J Pediatr 1992 Nov;121(5 Pt 1):710-714. [Medline: 1432418]
2. Lindenbaum J, Healton EB, Savage DG, Brust JC, Garrett TJ, Podell ER, et al. Neuropsychiatric disorders caused by cobalamin deficiency in the absence of anemia or macrocytosis. N Engl J Med 1988 Jun 30;318(26):1720-1728. [doi: 10.1056/NEJM198806303182604] [Medline: 3374544]
3. Almoallim H, Mehdawi FS, Cheikh MM, Al-Dhaheri F, Aqeel AM. Reversible vitamin B12 deficiency presenting with acute cementia, paraparesis, and normal hemoglobin. Case Rep Neurol Med 2016;2016:4301769 [FREE Full text] [doi: 10.1155/2016/4301769] [Medline: 28070432]
4. Tucker KL, Rich S, Rosenberg I, Jacques P, Dallal G, Wilson PW, et al. Plasma vitamin B-12 concentrations relate to intake source in the Framingham Offspring study. Am J Clin Nutr 2000 Feb;71(2):514-522 [FREE Full text] [Medline: 10648266]
5. Healton EB, Savage DG, Brust JC, Garrett TJ, Lindenbaum J. Neurologic aspects of cobalamin deficiency. Medicine (Baltimore) 1991 Jul;70(4):229-245. [Medline: 1648656]
6. Carmel R. How I treat cobalamin (vitamin B12) deficiency. Blood 2008 Sep 15;112(6):2214-2221 [FREE Full text] [doi: 10.1182/blood-2008-03-040253] [Medline: 18606874]
7. Naurath HJ, Joosten E, Riezler R, Stabler SP, Allen RH, Lindenbaum J. Effects of vitamin B12, folate, and vitamin B6 supplements in elderly people with normal serum vitamin concentrations. Lancet 1995 Jul 08;346(8967):85-89. [Medline: 7603218]
8. Komada Y, Narisawa H, Ueda F, Saito H, Sakaguchi H, Mitarai M, et al. Relationship between self-reported dietary nutrient intake and self-reported sleep duration among Japanese adults. Nutrients 2017 Feb 13;9(2):134 [FREE Full text] [doi: 10.3390/nu9020134] [Medline: 28208812]

9. Gougeon L, Payette H, Morais JA, Gaudreau P, Shatenstein B, Gray-Donald K. Intakes of folate, vitamin B6 and B12 and risk of depression in community-dwelling older adults: the Quebec Longitudinal Study on Nutrition and Aging. Eur J Clin Nutr 2016 Mar;70(3):380-385. [doi: 10.1038/ejcn.2015.202] [Medline: 26648330]
10. Almeida OP, Ford AH, Flicker L. Systematic review and meta-analysis of randomized placebo-controlled trials of folate and vitamin B12 for depression. Int Psychogeriatr 2015 May;27(5):727-737. [doi: 10.1017/S1041610215000046] [Medline: 25644193]
11. Vogiatzoglou A, Smith AD, Nurk E, Berstad P, Drevon CA, Ueland PM, et al. Dietary sources of vitamin B-12 and their association with plasma vitamin B-12 concentrations in the general population: the Hordaland Homocysteine Study. Am J Clin Nutr 2009 Apr;89(4):1078-1087 [FREE Full text] [doi: 10.3945/ajcn.2008.26598] [Medline: 19190073]
12. van Wijngaarden JP, Swart KMA, Enneman AW, Dhonukshe-Rutten RAM, van Dijk SC, Ham AC, et al. Effect of daily vitamin B-12 and folic acid supplementation on fracture incidence in elderly individuals with an elevated plasma homocysteine concentration: B-PROOF, a randomized controlled trial. Am J Clin Nutr 2014 Dec;100(6):1578-1586 [FREE Full text] [doi: 10.3945/ajcn.114.090043] [Medline: 25411293]
13. Yom-Tov E. Crowdsourced Health. Boston: MIT Press; 2016.
14. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection--harnessing the Web for public health surveillance. N Engl J Med 2009 May 21;360(21):2153-5, 2157 [FREE Full text] [doi: 10.1056/NEJMp0900702] [Medline: 19423867]
15. West R, White R, Horvitz E. From cookies to cooks: insights on dietary patterns via analysis of web usage log. 2013 Presented at: 22nd International Conference on World Wide Web; 2013 May 13-17; Rio de Janeiro, Brazil p. 1399-1410. [doi: 10.1145/2488388.2488510]
16. Trattner C, Parra D, Elsweiler D. Monitoring obesity prevalence in the United States through bookmarking activities in online food portals. PLoS One 2017;12(6):e0179144 [FREE Full text] [doi: 10.1371/journal.pone.0179144] [Medline: 28636665]
17. Abbar S, Mejova Y, Weber I. You tweet what you eat: studying food consumption through twitter. 2015 Presented at: ACM Conference on Human Factors in Computing Systems (CHI 15); 2015 Apr 18-23; Seoul, Republic of Korea p. 3197-3206.
18. Yom-Tov E, Gabrilovich E. Postmarket drug surveillance without trial costs: discovery of adverse drug reactions through large-scale analysis of web search queries. J Med Internet Res 2013 Jun 18;15(6):e124 [FREE Full text] [doi: 10.2196/jmir.2614] [Medline: 23778053]
19. Soldaini L, Yom-Tov E. Inferring individual attributes from search engine queries and auxiliary information. 2017 Presented at: 26th International Conference on World Wide Web; 2017 Apr 3-7; Perth, Australia p. 293-301.
20. Yom-Tov E, Borsa D, Hayward AC, McKendry RA, Cox IJ. Automatic identification of Web-based risk markers for health events. J Med Internet Res 2015 Jan 27;17(1):e29 [FREE Full text] [doi: 10.2196/jmir.4082] [Medline: 25626480]
21. Wikia Recipes. URL: http://recipes.wikia.com/wiki/Recipes_Wiki [accessed 2017-12-20] [WebCite Cache ID 6vqCxl51s]
22. Coppen A, Bolander-Gouaille C. Treatment of depression: time to consider folic acid and vitamin B12. J Psychopharmacol 2005 Jan;19(1):59-65. [doi: 10.1177/0269881105048899] [Medline: 15671130]
23. Mauro L, Martorana U, Cataldo P, Brancato G, Letizia G. Vitamin B12 in low back pain: a randomised, double-blind, placebo-controlled study. Eur Rev Med Pharmacol Sci 2000;4(3):53-58. [Medline: 11558625]
24. Yom-Tov E, White RW, Horvitz E. Seeking insights about cycling mood disorders via anonymized search logs. J Med Internet Res 2014 Feb 25;16(2):e65 [FREE Full text] [doi: 10.2196/jmir.2664] [Medline: 24568936]

Edited by G Eysenbach; submitted 08.08.17; peer-reviewed by C Trattner, C Basch, X Jiang, KNB Nor Aripin, MN Sadat; comments to author 19.10.17; revised version received 01.11.17; accepted 04.11.17; published 05.01.18.

Please cite as:
Giat E, Yom-Tov E
Evidence From Web-Based Dietary Search Patterns to the Role of B12 Deficiency in Non-Specific Chronic Pain: A Large-Scale Observational Study
J Med Internet Res 2018;20(1):e4
URL: http://www.jmir.org/2018/1/e4/
doi:10.2196/jmir.8667
PMID:29305340

Original Paper

Relationship Between State-Level Google Online Search Volume and Cancer Incidence in the United States: Retrospective Study

Charles A Phillips[1], MD; Allison Barz Leahy[1], MD; Yimei Li[1,2,3], PhD; Marilyn M Schapira[4,5], MD, MPH; L Charles Bailey[1,3,6], MD, PhD; Raina M Merchant[7,8], MD, MSHP

[1]Division of Oncology and Center for Childhood Cancer Research, The Children's Hospital of Philadelphia, University of Pennsylvania, Philadelphia, PA, United States

[2]Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania, Philadelphia, PA, United States

[3]Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[4]Department of Internal Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[5]Center for Health Equity Research and Promotion, Philadelphia Veterans Affairs Medical Center, Philadelphia, PA, United States

[6]Department of Biomedical and Health Informatics, The Children’s Hospital of Philadelphia, Philadelphia, PA, United States

[7]Penn Medicine Center for Digital Health, University of Pennsylvania, Philadelphia, PA, United States

[8]Department of Emergency Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

Corresponding Author:
Charles A Phillips, MD
Division of Oncology and Center for Childhood Cancer Research
The Children's Hospital of Philadelphia
University of Pennsylvania
Roberts Center for Pediatric Research, 11th Fl.
2716 South Street
Philadelphia, PA,
United States
Phone: 1 267 496 0161
Fax: 1 215 590 3296
Email: phillipsc2@email.chop.edu

Abstract

Background: In the United States, cancer is common, with high morbidity and mortality; cancer incidence varies between states. Online searches reflect public awareness, which could be driven by the underlying regional cancer epidemiology.

Objective: The objective of our study was to characterize the relationship between cancer incidence and online Google search volumes in the United States for 6 common cancers. A secondary objective was to evaluate the association of search activity with cancer-related public events and celebrity news coverage.

Methods: We performed a population-based, retrospective study of state-level cancer incidence from 2004 through 2013 reported by the Centers for Disease Control and Prevention for breast, prostate, colon, lung, and uterine cancers and leukemia compared to Google Trends (GT) relative search volume (RSV), a metric designed by Google to allow interest in search topics to be compared between regions. Participants included persons in the United States who searched for cancer terms on Google. The primary measures were the correlation between annual state-level cancer incidence and RSV as determined by Spearman correlation and linear regression with RSV and year as independent variables and cancer incidence as the dependent variable. Temporal associations between search activity and events raising public awareness such as cancer awareness months and cancer-related celebrity news were described.

Results: At the state level, RSV was significantly correlated to incidence for breast (r=.18, *P*=.001), prostate (r=–.27, *P*<.001), lung (r=.33, *P*<.001), and uterine cancers (r=.39, *P*<.001) and leukemia (r=.13, *P*=.003) but not colon cancer (r=–.02, *P*=.66). After adjusting for time, state-level RSV was positively correlated to cancer incidence for all cancers: breast (*P*<.001, 95% CI 0.06 to 0.19), prostate (*P*=.38, 95% CI –0.08 to 0.22), lung (*P*<.001, 95% CI 0.33 to 0.46), colon (*P*<.001, 95% CI 0.11 to 0.17), and uterine cancers (*P*<.001, 95% CI 0.07 to 0.12) and leukemia (*P*<.001, 95% CI 0.01 to 0.03). Temporal associations in GT were noted with breast cancer awareness month but not with other cancer awareness months and celebrity events.

Conclusions: Cancer incidence is correlated with online search volume at the state level. Search patterns were temporally associated with cancer awareness months and celebrity announcements. Online searches reflect public awareness. Advancing understanding of online search patterns could augment traditional epidemiologic surveillance, provide opportunities for targeted patient engagement, and allow public information campaigns to be evaluated in ways previously unable to be measured.

(J Med Internet Res 2018;20(1):e6) doi:10.2196/jmir.8870

KEYWORDS

Google; cancer; incidence; Internet; infodemiology

Introduction

Cancer is extremely common in the United States with over 1.5 million new diagnoses annually [1]. The 5 most common cancers in the United States are breast, prostate, colon, lung, and uterine [1]. The incidence for some of these cancers changes over time and varies between states [2,3]. Traditional epidemiologic methods from the Centers for Disease Control and Prevention (CDC) and National Cancer Institute (NCI) Surveillance, Epidemiology, and End Results (SEER) Program have a 2- to 4-year delay until incidence data are publicly reported [1]. While we acknowledge alternative cancer surveillance methods cannot replace traditional cancer surveillance and reporting methods, they have potential value if they are able to augment these gold standard methods in real time and offer information on public awareness for important cancer topics.

Approximately half of Americans report searching for cancer and health information online [4,5], and patients with cancer are increasingly seeking information on the Internet [6]. In addition to patients themselves, friends and family members are known to look up health information online for others [7]. Internet search data including Google Trends (GT) have been used to examine public interest in multiple health topics [8-14]. Google Flu Tracker (GFT), a program incorporating Google Correlate data but not GT, is perhaps the most prominent example of work comparing disease incidence to Google search data. In recent years, multiple studies have shown GFT can be inaccurate for a number of reasons including changes in underlying search rates, news coverage, changes in flu season severity, and errors in the algorithm itself [15-17]. The limitations with GFT must be considered in the design of any research focused on search data. In oncology, GT data have been used to examine multiple topics including seasonality of cancer interest [18], interest in cancer screening [19], efficacy of awareness campaigns [20,21], and effect of news coverage [22]. These events have been shown to drive cancer-related Internet activity and could serve to confound the relationship between cancer incidence and Internet search activity. One prior report demonstrated a correlation between Google search volume and cancer incidence and mortality [23] and did not adjust for changes in incidence over time or compare with drivers of online activity.

This study seeks to compare cancer incidence over time for 6 common cancers, as reported by surveillance registries, to GT data. This work is guided by the conceptual model that people with cancer and those in their immediate social networks are likely to use Google to seek information about cancer symptoms, diagnosis, therapies, side effects, and expected outcomes. Therefore, we hypothesized Google searches are reflective of state incidence patterns. Furthermore, we sought to evaluate previously reported sources of influence in Google data that could explain variability in our data.

Methods

Data Sources

This research involved free, publicly available, deidentified, online information from the CDC and GT for the years 2004 through 2013 and was deemed exempt from review by the Children's Hospital of Philadelphia institutional review board.

Cancer Incidence

National- and state-level annual cancer incidence was obtained from the CDC's website for each year of the study period for the 5 most common cancers in the United States, breast, prostate, lung and bronchus, colon and rectum, and corpus and uterine cancers, as well as for leukemia [2]. Leukemia was included as an example of a cancer that is present in both children and adults and could theoretically have a different search pattern compared to the solid tumors. Approximately 10% of leukemia cases are in children and adolescents compared to less than 1% of cases for the other cancers studied [2]. Cancer incidence data were collected for all 50 states and the District of Columbia except Nevada which did not have its incidence listed by the CDC for all study years. For secondary analysis, states and the District of Columbia were ranked by their cancer incidence from 1 to 50 for each cancer of interest.

Google Trends

Started in 2004, GT [24] is a free, publicly available, Internet-based application that allows the relative search frequency of different search terms or keywords to be compared to one another over time. It provides longitudinal data from 2004 through the present with the option to provide search data for specific geographic regions such as states or cities. GT presents search volume for a given term as a relative search volume (RSV) with a value between 0 and 100, with 100 being set as the most searched term in a given time period (weeks, months, or years) and other time periods assigned a proportionally lower number. For example, an RSV of 50 indicates half as many searches were performed in that time period compared to the time period with the highest volume of searches where RSV=100. An RSV of 0 indicates no searches were performed. RSV can either track relative interest in a region compared to itself or between that region and other regions. As detailed in the Statistical Analysis section, we used RSV to compare variation over time both within a state and between states. GT adjusts the RSV results for population size; results from populated areas are comparable to less populated areas.

We selected the search terms used in GT *a priori* using layman's terms for the common cancers. Our GT search terms were: "breast cancer," "prostate cancer," "colon cancer," "lung cancer," "uterine cancer," and "leukemia." Cancer search terms were entered into GT in 2 ways. First, for the primary analysis, each of the 6 cancer search terms was used individually to obtain the annual RSV for each state from 2004 through 2013. For example, if Kentucky were the state that searched for "breast cancer" the most in a given year, it would have a value of 100, while Kansas would have a value of 50 if it searched for "breast cancer" half as often as Kentucky in that year. Second, all cancer search terms were compared relative to each other for the United States as a whole for the study period, 2004 through 2013, to contextualize national trends for cancer type over time.

The secondary objective was to explore the impact of known drivers of Internet activity in our GT data. In this analysis, the RSV trends for each cancer were temporally compared to events previously reported to increase search activity, including cancer awareness months, celebrity events, and heavily covered news stories [19-22]. In this analysis, RSV was determined by setting 100 at the time period with greatest activity for the United States as a whole. Cancer awareness months for each of the 6 cancers of interest are October (breast), September (prostate), November (lung), March (colon), and September (uterine and leukemia). For other noticeable spikes in RSV, Google searches were performed in a 2-step process first using the cancer and date range to identify news stories for that timeframe: "lung cancer March 2010." If that search was not productive, a second search with the term "celebrity" was added: "lung cancer March 2010 celebrity."

Selection of Study Population

Our study focused on GT annual RSV for cancer by state for the years 2004 through 2013, selected because GT starts in 2004 and the most recent cancer incidence data published by the CDC is for the year 2013.

Statistical Analysis

Our outcome of interest was correlation between RSV and cancer incidence at the state level from 2004 through 2013. For all 6 cancers selected, we obtained annual RSV and cancer incidence for each state and the District of Columbia (up to 510 RSV and 510 incidence values per cancer). We used the Spearman correlation coefficient to examine the association between state-level GT RSV and state-level cancer incidence per 100,000 people for each year during the 10-year study period. Additionally, we examined the association of state-level RSV and state-level cancer incidence per 100,000 people using linear regression with RSV as the independent variable and state-level cancer incidence per 100,000 people as the dependent variable. For our linear regression, we included time in years as a continuous covariate to control for the fact that cancer incidence and RSV changed over the 10-year study period. RSV data were complete except for uterine cancer. Due to low search volumes for uterine cancer in sparsely populated states, RSV was not present in all states for all years, and we excluded 82 total missing values out of 510 potential observations (50 states and District of Columbia for 10 years).

In a secondary analysis, the aggregated (mean) cancer incidence for each state was obtained from the CDC for the years 2009 through 2013. A 5-year window was selected to reflect more recent cancer incidence and Internet search behaviors and use a shorter time period of aggregate data. The states were then ranked by aggregated cancer incidence and grouped into quartiles. The aggregated RSV for each state was then obtained from GT for the same 5-year period (2009 through 2013) and was similarly divided into quartiles. Quartiles were then compared to one another. For uterine cancer, RSV was not present for 10 states, and these states were excluded from analysis. Rank comparison for secondary analysis by rank quartiles was performed from 1 to 43 for uterine cancer and 1 to 50 for all other cancers.

Statistical analysis was conducted using Stata 14.2 (StataCorp LLC). A 2-sided *P*<.05 was considered statistically significant for all tests.

Results

Incidence and Google Search Volume

We examined state aggregate cancer incidence using CDC data and compared this with RSV in 2004 through 2013 (Multimedia Appendix 1). The median aggregate incidence per 100,000 people for each cancer from 2004 through 2013 was as follows: breast 123.5, prostate 142.3, lung 66.7, colon 44.8, and uterine cancer 25.1 and leukemia 13.5. Regarding cancer incidence, prostate and colon cancer decreased over the study period (Figure 1). Cancer incidence for the other cancers studied was relatively constant. The median aggregate RSV over the study period for each cancer was breast 66.7, prostate 63.2, lung 62.5, colon 66.9, and uterine cancers 53.2 and leukemia 70.5. These RSV represent the searches for each cancer term compared between states. In addition, we compared the RSV of the cancer terms to each other during the 2004 through 2013 study period for the United States as a whole, with breast cancer being the most searched (Figure 1).

Correlation and Regression Comparing State-Level Cancer Incidence and Google Trends Relative Search Volume

Using the Spearman correlation, state-level RSV was significantly correlated to state-level cancer incidence for breast (*r*=.18, *P*=.001), prostate (*r*=–.27, *P*<.001), lung (*r*=.33, *P*<.001), and uterine cancers (*r*=.39, *P*<.001) and leukemia (*r*=.13, *P*=.003) but not for colon cancer (*r*=–.02, *P*=.66) (Table 1). Linear regression demonstrated consistent results for the positive and negative correlations seen using the Spearman correlation. After adding time to our regression to account for changes in cancer incidence and RSV, the coefficients for all cancers were positive with a range of .02 (95% CI 0.01 to 0.03) for leukemia to .39 (95% CI 0.33 to 0.46) for lung cancer (Table 1). Prostate cancer was the only cancer not to have a statistically significant positive association (*P*=.38, 95% CI –0.08 to 0.22). The coefficients of determination ranged from a low of .03 for leukemia to a high of .56 for colon cancer.

Figure 1. Relative incidence versus Google relative search volume for common cancers 2004-2013.

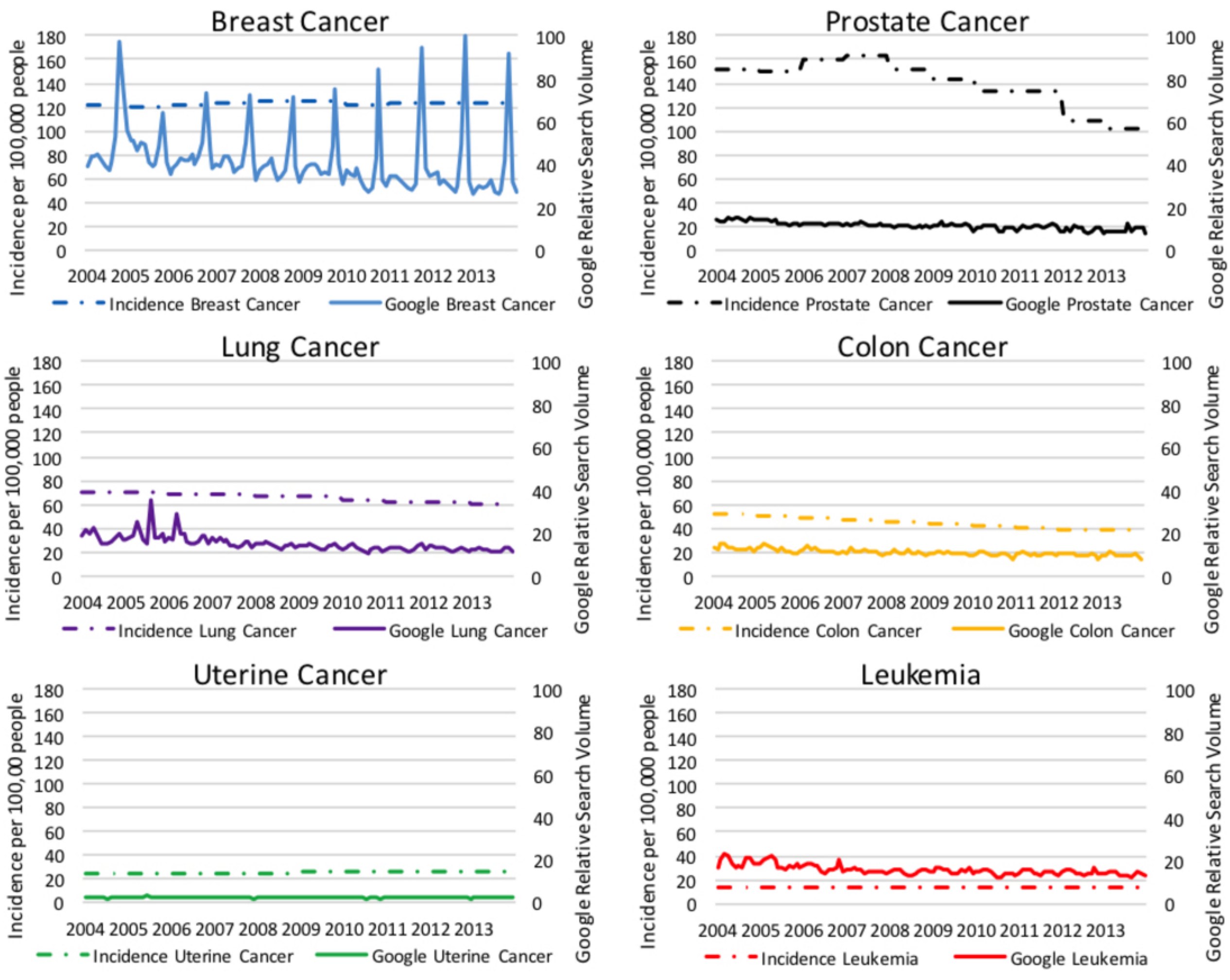


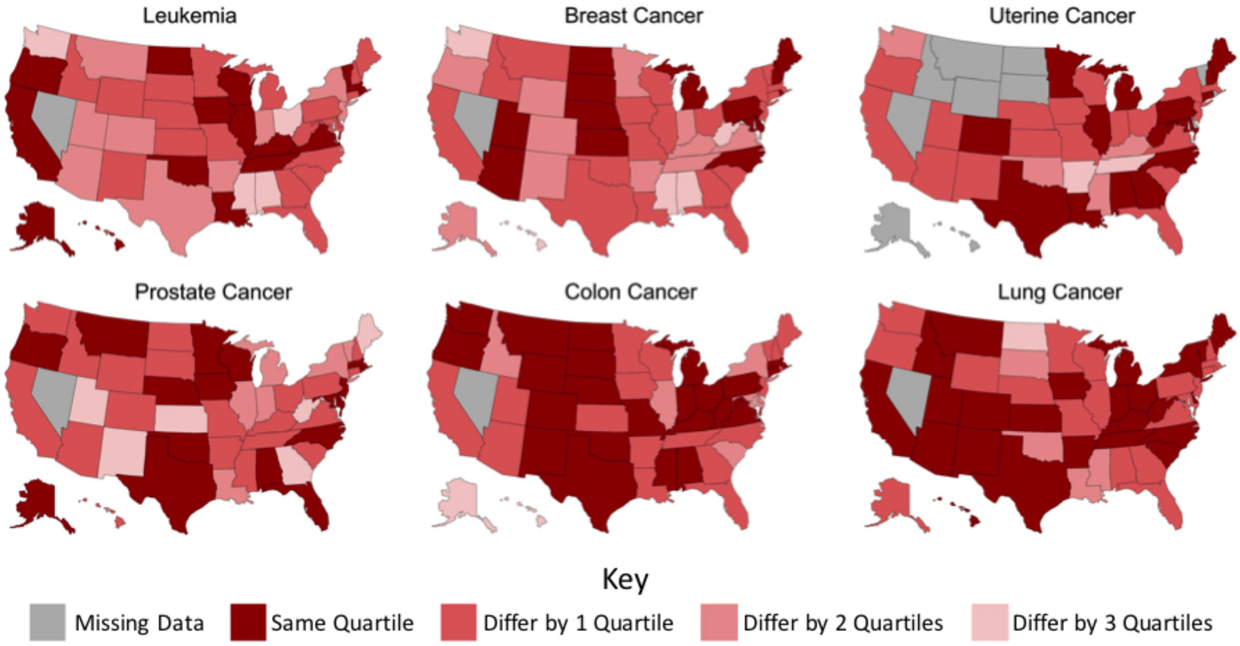
Table 1. Correlation and linear regression for state-level Google relative search volume and cancer incidence from 2004 through 2013.

CDC[a] cancer term	Google cancer term	Correlation test		Linear regression with year included		
		Spearman r_s	Spearman P value	Coefficient (95% CI)	P value	R^2
Breast	Breast cancer	.18	.001	.12 (0.06 to 0.19)	<.001	.05
Prostate	Prostate cancer	–.27	<.001	.07 (–0.08 to 0.22)	.38	.47
Lung and bronchus	Lung cancer	.33	<.001	.39 (0.33 to 0.46)	<.001	.28
Colon and rectum	Colon cancer	–.02	.67	.14 (0.11 to 0.17)	<.001	.56
Corpus and uterus, NOS[b]	Uterine cancer	.40	<.001	.09 (0.07 to 0.12)	<.001	.16
Leukemias	Leukemia	.13	.003	.02 (0.01 to 0.03)	.001	.03

[a] CDC: Centers for Disease Control and Prevention.

[b] NOS: not otherwise specified.

Figure 2. State cancer incidence rank and Google relative search volume (RSV) rank for common cancers by quartile 2009-2013. States whose rank-based quartile was the same for both cancer incidence and RSV are shown in dark red. Progressively lighter shades indicate greater difference in rank of cancer incidence and RSV (quartile difference ranged from 0-3). States shown in gray had missing data and were excluded.



Google Trends. Interest by subregion. https://www.google.com/trends/explore?date=2009-01-01%202011-12-31&geo=US&q=lung%20cancer. Accessed 1/18/2017

Comparing Cancer Incidence Rank and RSV Rank by Quartile

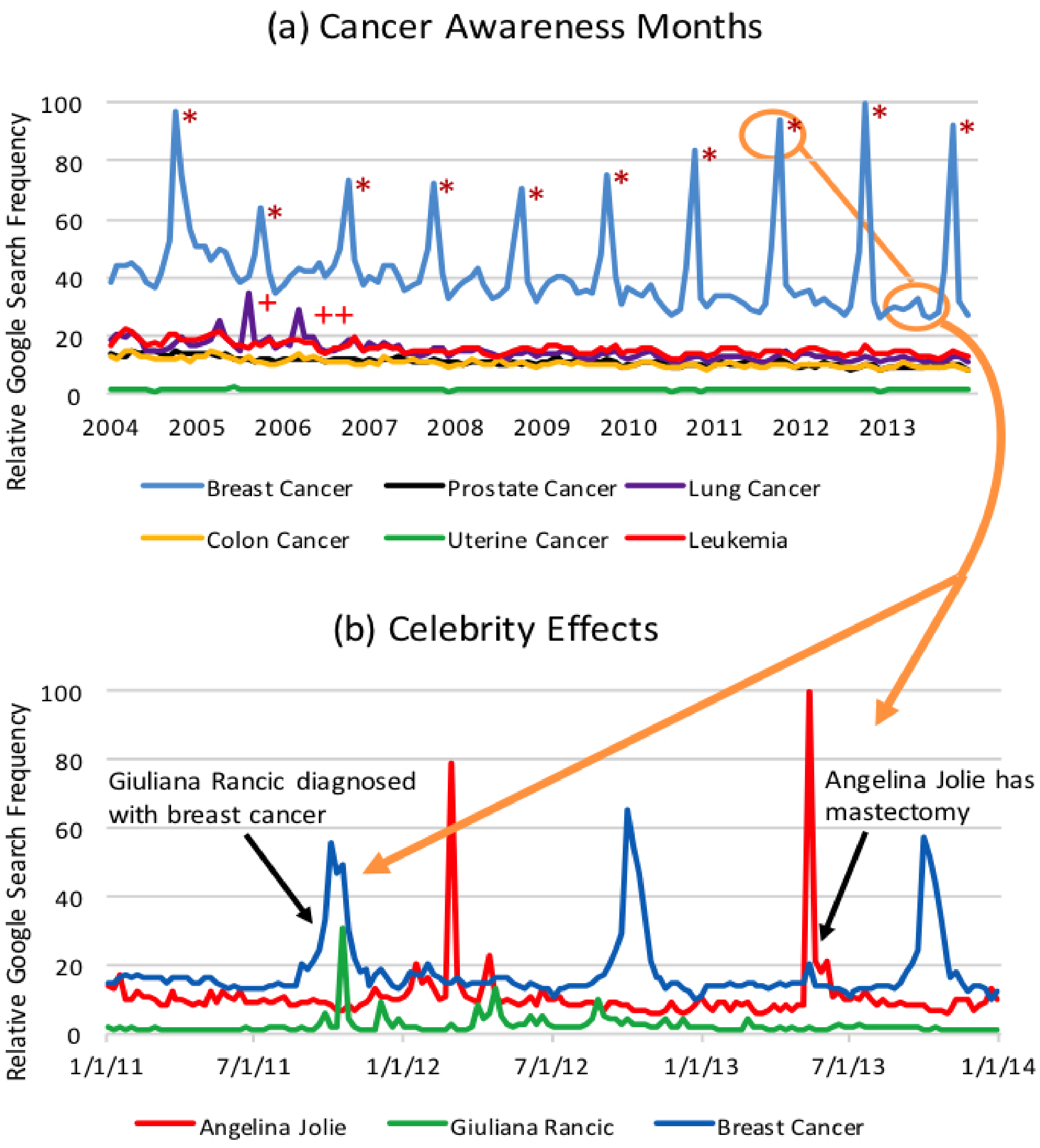
The states were ranked in order based on their cancer incidence and RSV from 2009 through 2013 and grouped into quartiles. Figure 2 depicts a map of the United States highlighting the degree of agreement by quartile for a state's average cancer incidence and average RSV. When grouped by quartile, some cancers demonstrate a higher agreement of state cancer incidence rank and state RSV rank than others.

Cancer Awareness Months and News Events

Search spikes were common for breast cancer and less so for the other cancers. Breast cancer has a large increase in searches during the month of October each year (Figure 1). These annual spikes represent increases in search volume that are more than double the baseline search volume for breast cancer and are temporally associated with breast cancer awareness month. Excluding breast cancer awareness month from our time-adjusted regression raised the coefficient of determination from .05 to .07. No other cancer had a spike in activity during its awareness month.

Lung cancer was the only other cancer to have at least 1 RSV spike that was double its baseline RSV. The 2 RSV spikes for lung cancer occurred in August 2005 and March 2006, which were temporally associated with the deaths of public figures (Peter Jennings and Dana Reeve) from lung cancer (Figure 3). For breast cancer, spikes in RSV were temporally noted with Angelina Jolie's mastectomy and Guiliana Rancic's public announcement of breast cancer (Figure 3).

Figure 3. Temporal relationship of Google Trend data to public events. (A) Trends in relative search volume (RSV) for all 6 cancers of interest from 2004-2013. *Breast cancer awareness month. +Death of Peter Jennings. ++Death of Dana Reeve. (B) Independent RSV trends for Angelina Jolie, Guiliana Rancic, and breast cancer.



Discussion

Principal Findings

In this study, we examined the association of Google search activity with cancer incidence over time across the United States. Our results demonstrate a significant association between Google search activity and incidence of 5 of the 6 common cancers at the state level. The strength of association between RSV and cancer incidence varied among the cancers studied. Conclusions drawn from online search volume about one type of cancer may not be able to be generalized to other types of cancer. We see a similar pattern of limited generalizability in studies of the relationship of online search activity to disease incidence for subtypes of influenza where searches for H1N1 were different compared to other types of influenza [15-16] and, in a prior study within oncology, incidence positively correlated to Google search volume for 5 of 8 cancers studied [23]. Similarly, we found a positive correlation for 4 of 6 cancers, and the strength of positive correlation for all 6 improved when time was added to a regression model, with only prostate cancer failing to reach significance.

Incorporating time into the model had the greatest impact for colon and prostate cancer due to their declining incidence during the study period; both changed from a negative Spearman correlation coefficient to a positive regression coefficient (Table 1). Because the RSVs are set to 100 for each year in the analysis, a declining underlying population incidence effectively changes what 100 represents in a given year. Caution should be taken when interpreting online data for diseases with an incidence that changes over time.

Other lessons applicable to Internet surveillance research involve the cautionary tale of the GFT. After initial success, it proved to overestimate flu incidence as a result of mismatching correlated terms, different media coverage levels between flu seasons, and lack of algorithm transparency [17]. Our work differs in that it does not use correlated search terms, and cancer has a different online search profile compared to infectious diseases. For example, cancer is typically not searched for on a seasonal basis, and media coverage of cancer is likely to be different than the flu. We attempt to account for the media coverage by examining searches relative to awareness months and explicitly searching for news stories when the RSV data showed an unexpected rise.

Temporal associations are present for news stories and popular culture. The breast cancer RSV curve has spikes in activity that are temporally related to reports of Angelina Jolie’s mastectomy and Guiliana Rancic’s public announcement of breast cancer (Figure 3). The “Jolie Effect” has been described in increases in breast cancer susceptibility gene (BRCA) testing following her public disclosure [25] and more websites addressing common themes regarding care for patients with BRCA mutations after her public announcement [26]. Knowing the news pulse for specific stories such as Ms. Jolie’s mastectomy in response to her BRCA status offers opportunities for targeted medical messaging by the public health community that overlaps with an increase in public interest in that topic.

Other celebrity events temporarily associated with RSV spikes include the deaths of Peter Jennings and Dana Reeve from lung cancer. If search data are to be considered for the purpose of surveillance, current events and drivers of Internet activity must be taken into account, as these drivers may obscure the relationship between searches performed in response to direct impact on individuals and those driven by news or public information. Further work is needed to clarify these potential confounders of the relationship between cancer incidence and search activity and improve the utility of Google search data.

Despite factors other than cancer incidence driving searches, we found online search activity mirrors cancer incidence. Cancer clusters with unusually high incidence have been reported [27], and online adjunctive surveillance may have been useful in detecting these hot spots. Online search activity is less likely to be relevant in trending the national incidence, which has well-established surveillance and reporting methods. With appropriate transparency in trend algorithms, further work, and appropriate input from the scientific community, meaningful public health initiatives and adjunctive cancer surveillance methods could be achieved.

In addition to detecting signals about cancer incidence, search data are exquisitely good at reflecting people’s interest at the population level. Our data add to the literature supporting news coverage, and cancer awareness campaigns can register with the general public and drive online activity. It remains unclear why breast cancer is the only cancer studied that demonstrated a significant increase in its RSV during its cancer awareness month. The granularity of RSV data allows for assessment of the impact of public health campaigns and public awareness at the national, state, and metropolitan area levels. The most practical current application for online surveillance may be assessing changes in public engagement after an event or campaign to increase public knowledge about a cancer topic.

Online data provide information about which cancers could be the best targets for digital outreach. For example, prostate cancer was the only cancer studied without a significant association between RSV and incidence and may be a poor choice for online interventions. One explanation for the lack of significant association could be that the population with prostate cancer tends to be elderly men. According to the US Census Bureau, in 2010 only 55% of people aged 65 years and older had Internet access in their homes [28]. It is possible that association with RSV for prostate cancer will become significant as the current population integrated into Internet use ages. Understanding the relevance of online searches for a given cancer could inform patient-centered approaches to distribute information for many aspects of cancer care including trial recruitment, screening practices, and care options.

Finally, establishing a link between online search activity and cancer incidence is of use to those interested in mining the Internet and social mediome for medically relevant information. Internet searches provide data that indicate what people want to know and when they want to know it. Linking cancer search volume to incidence provides validity to work examining correlations between cancer terms and other aspects of cancer care such as treatment side effects.

Limitations

This study has some limitations. It did not include searches done through alternative search engines such as Yahoo or Bing. The algorithm employed by Google to determine the RSV was not published and could contain systematic errors. GT reported data at the state level and for major metropolitan areas. It was less well suited for rural areas and uncommon topics. Additionally, GT cannot link to a specific user, and data were only available at the population level. We cannot therefore control for confounders that may impact search activity and cancer incidence such as race, ethnicity, smoking status, socioeconomic status, and level of education.

Conclusion

Our 3 key findings are cancer incidence is correlated with Google search volume at the state level, different cancers demonstrate unique Google search patterns, and search patterns are influenced by public events such as cancer awareness months and news coverage of celebrity experiences with cancer. Online searches reflect public awareness, and advancing understanding of online search patterns could augment traditional epidemiologic surveillance, provide opportunities for targeted patient engagement, and allow public information campaigns to be evaluated in ways previously unable to be measured.

Acknowledgments

Research reported in this publication was supported by a grant to CAP and ABL (T32HD060550) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health. RMM receives funding from the Robert Wood Johnson Foundation (Pioneer Award 72695).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Average incidence per 100,000 people and average relative search volume by state from 2004-2013.

[PDF File (Adobe PDF File), 33KB - jmir_v20i1e6_app1.pdf]

References

1. Siegel R, Miller K, Jemal A. Cancer statistics, 2017. CA Cancer J Clin 2017 Jan;67(1):7-30 [FREE Full text] [doi: 10.3322/caac.21387] [Medline: 28055103]
2. United States cancer statistics: 1999-2014 incidence and mortality Web-based report. Atlanta: Centers for Disease Control and Prevention; 2017.
3. Jemal A, Ward EM, Johnson CJ, Cronin KA, Ma J, Ryerson B, et al. Annual report to the nation on the status of cancer, 1975-2014, featuring survival. J Natl Cancer Inst 2017 Dec 01;109(9):1 [FREE Full text] [doi: 10.1093/jnci/djx030] [Medline: 28376154]
4. Arora NK, Hesse BW, Rimer BK, Viswanath K, Clayman ML, Croyle RT. Frustrated and confused: the American public rates its cancer-related information-seeking experiences. J Gen Intern Med 2008 Mar;23(3):223-228 [FREE Full text] [doi: 10.1007/s11606-007-0406-y] [Medline: 17922166]
5. Diaz JA, Griffith RA, Ng JJ, Reinert SE, Friedmann PD, Moulton AW. Patients' use of the Internet for medical information. J Gen Intern Med 2002 Mar;17(3):180-185 [FREE Full text] [Medline: 11929503]
6. Kowalski C, Kahana E, Kuhr K, Ansmann L, Pfaff H. Changes over time in the utilization of disease-related Internet information in newly diagnosed breast cancer patients 2007 to 2013. J Med Internet Res 2014 Aug 26;16(8):e195 [FREE Full text] [doi: 10.2196/jmir.3289] [Medline: 25158744]
7. Ayantunde AA, Welch NT, Parsons SL. A survey of patient satisfaction and use of the Internet for health information. Int J Clin Pract 2007 Mar;61(3):458-462. [doi: 10.1111/j.1742-1241.2006.01094.x] [Medline: 17313614]
8. Bragazzi NL, Barberis I, Rosselli R, Gianfredi V, Nucci D, Moretti M, et al. How often people google for vaccination: qualitative and quantitative insights from a systematic search of the web-based activities using Google Trends. Hum Vaccin Immunother 2017 Feb;13(2):464-469. [doi: 10.1080/21645515.2017.1264742] [Medline: 27983896]
9. Cho S, Sohn CH, Jo MW, Shin S, Lee JH, Ryoo SM, et al. Correlation between national influenza surveillance data and google trends in South Korea. PLoS One 2013;8(12):e81422 [FREE Full text] [doi: 10.1371/journal.pone.0081422] [Medline: 24339927]
10. Choi J, Cho Y, Shim E, Woo H. Web-based infectious disease surveillance systems and public health perspectives: a systematic review. BMC Public Health 2016 Dec 08;16(1):1238 [FREE Full text] [doi: 10.1186/s12889-016-3893-0] [Medline: 27931204]
11. Gamma A, Schleifer R, Weinmann W, Buadze A, Liebrenz M. Could Google Trends be used to predict methamphetamine-related crime? An analysis of search volume data in Switzerland, Germany, and Austria. PLoS One 2016;11(11):e0166566 [FREE Full text] [doi: 10.1371/journal.pone.0166566] [Medline: 27902717]
12. 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]
13. Mendonça MD, Caetano A, Viana-Baptista M. Dr Google will see you now—time trends in online searches on headache. Cephalalgia 2016 Nov 28:1. [doi: 10.1177/0333102416681572] [Medline: 27895210]
14. Polgreen PM, Chen Y, Pennock DM, Nelson FD. Using internet searches for influenza surveillance. Clin Infect Dis 2008 Dec 1;47(11):1443-1448 [FREE Full text] [doi: 10.1086/593098] [Medline: 18954267]
15. Butler D. When Google got flu wrong. Nature 2013 Feb 14;494(7436):155-156. [doi: 10.1038/494155a] [Medline: 23407515]
16. 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]
17. Lazer D, Kennedy R, King G, Vespignani A. Big data. The parable of Google Flu: traps in big data analysis. Science 2014 Mar 14;343(6176):1203-1205. [doi: 10.1126/science.1248506] [Medline: 24626916]
18. Zhang Z, Zheng X, Zeng DD, Leischow SJ. Information seeking regarding tobacco and lung cancer: effects of seasonality. PLoS One 2015;10(3):e0117938 [FREE Full text] [doi: 10.1371/journal.pone.0117938] [Medline: 25781020]

19. Schootman M, Toor A, Cavazos-Rehg P, Jeffe DB, McQueen A, Eberth J, et al. The utility of Google Trends data to examine interest in cancer screening. BMJ Open 2015;5(6):e006678 [FREE Full text] [doi: 10.1136/bmjopen-2014-006678] [Medline: 26056120]
20. Kantor J. Association of Google search volume index peaks for skin cancer with skin cancer awareness month. JAMA Dermatol 2016 Jan;152(1):113. [doi: 10.1001/jamadermatol.2015.3814] [Medline: 26762257]
21. Murray G, O'Rourke C, Hogan J, Fenton JE. Detecting internet search activity for mouth cancer in Ireland. Br J Oral Maxillofac Surg 2016 Feb;54(2):163-165. [doi: 10.1016/j.bjoms.2015.12.005] [Medline: 26774361]
22. Fazeli DS, Carlos RC, Hall KS, Dalton VK. Novel data sources for women's health research: mapping breast screening online information seeking through Google trends. Acad Radiol 2014 Sep;21(9):1172-1176 [FREE Full text] [doi: 10.1016/j.acra.2014.05.005] [Medline: 24998689]
23. Wehner MR, Nead KT, Linos E. Correlation among cancer incidence and mortality rates and Internet searches in the United States. JAMA Dermatol 2017 Jun 28;153(9):911-914. [doi: 10.1001/jamadermatol.2017.1870] [Medline: 28658470]
24. Google Trends. URL: https://trends.google.com/trends/ [accessed 2017-12-07] [WebCite Cache ID 6sxlCSCyi]
25. Roberts MC, Dusetzina SB. The effect of a celebrity health disclosure on demand for health care: trends in BRCA testing and subsequent health services use. J Community Genet 2017 Apr;8(2):141-146 [FREE Full text] [doi: 10.1007/s12687-017-0295-7] [Medline: 28299592]
26. Dean M. Celebrity health announcements and online health information seeking: an analysis of Angelina Jolie's preventative health decision. Health Commun 2016;31(6):752-761. [doi: 10.1080/10410236.2014.995866] [Medline: 26574936]
27. Mokdad AH, Dwyer-Lindgren L, Fitzmaurice C, Stubbs RW, Bertozzi-Villa A, Morozoff C, et al. Trends and patterns of disparities in cancer mortality among US counties, 1980-2014. JAMA 2017 Dec 24;317(4):388-406. [doi: 10.1001/jama.2016.20324] [Medline: 28118455]
28. Computer and Internet use in the United States.: US Census Bureau; 2013. URL: https://www.census.gov/content/dam/Census/library/publications/2013/demo/p20-569.pdf [accessed 2017-12-07] [WebCite Cache ID 6vXrPHwGM]

Abbreviations

BRCA: breast cancer susceptibility gene
CDC: Centers for Disease Control and Prevention
GFT: Google Flu Tracker
GT: Google Trends
NCI: National Cancer Institute
RSV: relative search volume
SEER: Surveillance, Epidemiology, and End Results

Edited by G Eysenbach; submitted 30.08.17; peer-reviewed by Z Yan, N Bragazzi, M Martini; comments to author 04.10.17; revised version received 18.10.17; accepted 29.10.17; published 08.01.18.

Please cite as:
Phillips CA, Barz Leahy A, Li Y, Schapira MM, Bailey LC, Merchant RM
Relationship Between State-Level Google Online Search Volume and Cancer Incidence in the United States: Retrospective Study
J Med Internet Res 2018;20(1):e6
URL: http://www.jmir.org/2018/1/e6/
doi:10.2196/jmir.8870
PMID:29311051

Original Paper

Exploring Entertainment Medicine and Professionalization of Self-Care: Interview Study Among Doctors on the Potential Effects of Digital Self-Tracking

Katleen Gabriels[1*], PhD; Tania Moerenhout[2,3*], MPhil, MD

[1] Philosophy & Ethics, Department of Industrial Engineering & Innovation Sciences, Eindhoven University of Technology, Eindhoven, Netherlands
[2] Ethics, Autonomy and Responsibility in Health Care, Department of Family Medicine and Primary Health Care, Ghent University, Ghent, Belgium
[3] Department of Philosophy and Moral Sciences, Ghent University, Ghent, Belgium
[*] all authors contributed equally

Corresponding Author:
Katleen Gabriels, PhD
Philosophy & Ethics, Department of Industrial Engineering & Innovation Sciences
Eindhoven University of Technology
Office: IPO 1.07
Postbus 513
Eindhoven, 5600 MB
Netherlands
Phone: 31 40 247 5161
Fax: 31 40 244 46 02
Email: k.gabriels@tue.nl

Abstract

Background: Nowadays, digital self-tracking devices offer a plethora of possibilities to both healthy and chronically ill users who want to closely examine their body. This study suggests that self-tracking in a private setting will lead to shifting understandings in professional care. To provide more insight into these shifts, this paper seeks to lay bare the promises and challenges of self-tracking while staying close to the everyday professional experience of the physician.

Objective: The aim of this study was to (1) offer an analysis of how medical doctors evaluate self-tracking methods in their practice and (2) explore the anticipated shifts that digital self-care will bring about in relation to our findings and those of other studies.

Methods: A total of 12 in-depth semistructured interviews with general practitioners (GPs) and cardiologists were conducted in Flanders, Belgium, from November 2015 to November 2016. Thematic analysis was applied to examine the transcripts in an iterative process.

Results: Four major themes arose in our body of data: (1) the patient as health manager, (2) health obsession and medicalization, (3) information management, and (4) shifting roles of the doctors and impact on the health care organization. Our research findings show a nuanced understanding of the potentials and pitfalls of different forms of self-tracking. The necessity of contextualization of self-tracking data and a professionalization of self-care through digital devices come to the fore as important overarching concepts.

Conclusions: This interview study with Belgian doctors examines the potentials and challenges of self-monitoring while focusing on the everyday professional experience of the physician. The dialogue between our dataset and the existing literature affords a fine-grained image of digital self-care and its current meaning in a medical-professional landscape.

(J Med Internet Res 2018;20(1):e10) doi:10.2196/jmir.8040

KEYWORDS

mobile applications; wearable electronic devices; self-recorded health data; self care; quantified self; qualitative research

Introduction

Today, individuals are offered a multitude of possibilities to track and so manage their personal health. Wearables such as Jawbone and Fitbit allow people to monitor bodily processes and activities. Self-tracking through wearable and otherwise mobile computing has become known as the *quantified self* [1]. Quantified self technologies—usually small computers that record data—provide the individual with detailed information including sleeping habits and calories burnt. This knowledge can lead to changes in self-understanding [2].

Although self-care has been an established practice for several years (eg, the use of home blood-pressure monitors), digital self-care by means of mobile computing—so-called mobile health (mHealth)—is still in its introductory stages. As an initial step to understanding the opportunities, as well as ethical challenges that mHealth tools present, we examine how both classic and digital self-tracking methods are incorporated into daily health care.

Classic self-tracking methods provide support to enhance self-care: “the care of oneself without medical, professional, or other assistance or oversight” [3]. The concept is complex, encompassing various aspects, and changing over time. Here is a more comprehensive definition [3]:

> *Self-care is deliberate care performed throughout life; by individuals to themselves and to others; to promote health or improve both general health and mental health, and cope with illness or disability; and in collaboration with healthcare professionals or performed separately. Self-care also includes social support and provides the continuity of care necessary to maintain wellbeing.*

When focusing on the management of chronic disease, the term “self-management” is appropriate, defined as the “patient’s ability to manage the symptoms, treatment, physical and psychosocial consequences, and lifestyle changes inherent in living with a chronic condition” [4]. The application of self-analytic methods is not new. Indeed, physicians have been recommending self-tracking for some time, both for self-management (eg, diabetes) and professional care (eg, cardiac monitoring). Before apps and wearable technologies became widespread, self-tracking for self-care relied on more simple technologies such as thermometers and bathroom scales.

Researchers investigating the application of digital self-tracking to professional care vary from the enthusiastic, focusing on the technology’s revolutionary promise [5], to the skeptical, focusing on the risks [6]; with any number taking moderate standpoints in between [7]. Health apps and devices promise to discover medical problems earlier and also serve to coach the user toward behavioral changes: for example, more exercise. Self-tracking technologies are appropriate both for healthy persons (preventive care) and for the ill (interventionist or therapeutic use)—in which case, the technology assists in keeping a close watch on the condition, noting any anomalies. In both cases, the data can be shared with a professional health care provider.

This paper has a twofold objective: first, to analyze how medical doctors evaluate both classic and digital self-tracking on an everyday clinical basis. Toward this end, we conducted semistructured interviews with Belgian general practitioners (GPs; n=7) and cardiologists (n=5) to gain insight into their shifting understandings of these technologies. How do they incorporate self-tracking in their daily practice? What ethical issues do they encounter? What are their expectations for the future of digital self-tracking? Second, we sought to explore the transitions that digital self-tracking will bring about in self-care and professional health care by relating existing literature to care providers’ everyday experience.

Throughout this paper, classic is distinguished from digital self-tracking based on two main characteristics: data sharing and who takes the initiative to track. First, in the case of classic self-tracking, data are usually collected for private use only, whereas sharing is generally limited to care providers and mostly paper-based. The care provider typically initiates the request to self-track and, at least in some cases, provides the device or gives recommendations on which device to use and how to track (*cf.* “pushed” self-tracking, ie, the incentive to self-track comes from an external actor [8]). Second, digital self-tracking is most often Internet-connected, mobile phone–based, and designed from the start for data sharing. Members of the quantified self movement share their data with each other through social media (*cf.* “communal” self-tracking, ie, the practice of sharing data with a community of trackers [8]). It is usually the patient or healthy user who initiates the tracking. In some cases, the app takes over the care provider’s role. Coaching apps are designed to motivate behavioral change. Collected data can easily be shared with care providers and become part of patients’ electronic health record (EHR). The distinction of classical from digital self-tracking is not clear-cut. Rather, it is situated on a continuum, which facilitates comparison.

Methods

Sample Selection

This study used purposive, convenience-based sampling to select care providers, aiming at maximum variation in expertise, gender, years of experience, geographical location, and type of practice. We selected doctors with relevant experience with classic self-tracking. The heterogeneous sample served to develop, first, a broader insight into classic and digital self-tracking in everyday health care and, second, a range of perspectives by viewing the matter through various lenses. One author (TM), who works part-time as a GP, facilitated the recruiting process. This helped us gain access to this specific group of experts who are often hard to reach. In qualitative studies such as ethnographies, it is not uncommon to conduct research in the group one already has access to or in the group one already is a part of [9].

We arrived at a point of saturation after seven interviews [10,11]. Our participants shared similar views on most of the topics we talked about throughout the interviews and raised similar themes, suggesting that we reached a point where we could reasonably expect to not collect any further contrasting results, even within our small sample. Despite their experience with classic self-tracking, the GPs had little experience with digital self-tracking, and this was not often brought up by patients. Digital self-care is not yet an integrated part of their practice.

We subsequently interviewed cardiologists selected on the basis of experience with digital self-tracking and telemonitoring (in their case, remote tracking of patients with an implanted pacemaker or defibrillator). We arrived at a point of saturation after five interviews.

The Ethical Commission of Human Sciences (Vrije Universiteit Brussel) approved the study. All respondents are Belgian and Dutch-speaking. They gave their consent to participate and were informed about the researchers and research context. The respondents signed an informed consent and consented to a digital audio recording (see Multimedia Appendix 1 for the informed consent form). All interviews were conducted face-to-face and took place in the office of the physicians.

Interview Guide

The interviews were semistructured: the interviewer started from an interview guide with a set of predetermined topics but equally left scope for extra questions or topics (see Multimedia Appendix 2). The interview guide for GPs was divided into two sections. We started with questions aimed at understanding the current situation: how do the providers incorporate classic self-tracking methods in their everyday practice? We focused on blood pressure meters, oximeters, and glucose meters as concrete examples. The questions were organized into five topics of which the last was the most elaborately discussed topic:

1. Frequency of use
2. Initiative to self-track
3. Choice and quality of device
4. Data interpretation
5. Positive and negative results or effects of self-tracking

The second part of the interview focused on digital self-tracking. The GPs were presented with five examples of apps. The questions were organized in the same manner as the first part of the interview, with three additional topics:

1. Differences between classic and digital self-tracking
2. Data sharing
3. Inclusion (health disparities)

We ended the interview with a peek into the future, asking them to imagine their practice 10 years from now and the evolution they expected in self-tracking and the health care organization.

The interview guide developed for cardiologists was similar to the first one, except for minor changes that made it more relevant to the cardiology practice, for example, focusing on blood pressure and heart rate apps.

Data Analysis

All interviews were transcribed ad verbatim by one of the researchers (KG transcribed 7 interviews and TM 5). We performed a thematic analysis carried out in three phases. The *first phase* involved labeling and tagging the data and assigning codes to the text by marking words and phrases in Word (Microsoft) [12]. The first interview was coded by both researchers, followed by a thorough discussion of the developed codes until consensus was reached. One researcher (TM) continued to code all interviews in an iterative process. The second researcher (KG) independently coded three more interviews to ensure validity and reliability of the codes. This was again followed by an elaborate discussion of code definition until consensus was reached. Both authors were also particularly attentive to deviant cases in the interviews. A coding scheme was subsequently developed that first derived *deductive* codes based on the literature review and research questions. Second, the coding scheme also consisted of *inductive* codes, emerging directly from the interviews (see Multimedia Appendix 3 for the coding scheme).

The *second phase* consisted of an in-depth search for deeper themes in the codes. In doing so, the data were recoded thematically, and patterns were looked for between the codes to identify and generate core and deeper themes [13].

The *third phase* involved the move from themes to theory: we linked the themes with academic literature on self-care and embedded them in broader frameworks. The selected quotes were translated from Dutch to English by the researchers. One researcher (KG) also has a master’s degree in English (linguistics and literature), which helped assure an accurate translation.

To enhance quality, reliability, and validity of the research, we also included member validation [14]. Three cardiologists and four GPs of our sample reviewed a version of this manuscript and its research findings. We invited them to comment upon our interpretations [15]. All of them agreed with our study findings. Some participants gave a number of constructive suggestions; we subsequently took their remarks into consideration and included them in the manuscript.

Interviews

All interviews were conducted from November 2015 to November 2016. One researcher (TM) carried out eight interviews (six GPs and two cardiologists); the second researcher (KG) conducted four interviews (one GP and three cardiologists). The shortest interview was 38 min (transcript of 5400 words), and the longest was 105 min (transcript of 14,200 words). As described above, we only selected GPs who had extensive experience with classic self-tracking in their daily practice and aimed for a heterogeneous sample. One GP has specific experience as an information technology (IT) expert for a local GP network. Cardiologists 11 and 12 are also engaged in research projects on this subject. A schematic overview of the sample can be found in Table 1.

Table 1. Characteristics of the sample.

Number	Occupation	Gender	Age range (years)[a]	Years of experience	Type of practice
1	General practitioner	Female	26-30	2	Duo
2	General practitioner	Female	31-35	4	Group
3	General practitioner	Female	36-40	9	Solo
4	General practitioner	Male	51-55	7	Solo
5	General practitioner	Female	56-60	30	Solo for 26 years, now group
6	General practitioner	Female	41-45	15	Health community center
7	General practitioner	Male	66-70	40	Duo
8	Cardiologist	Female	46-50	15	Hospital
9	Cardiologist	Male	56-60	25	Hospital
10	Cardiologist	Male	36-40	4	Hospital
11	Cardiologist	Female	26-30	2 (assistant)	Hospital and research
12	Cardiologist	Male	56-60	20+	Hospital and research

[a] For the sake of confidentiality, we opted for providing the age range instead of the year of birth.

Belgian Situation

Given that the interviews are conducted with Belgian health care providers, we concisely describe the Belgian health care context. Belgium is an active welfare state with an extensive form of social security, covered by social contributions based on income [16]. One of the six sectors of the social security system is a compulsory health insurance with a broad benefits package covering almost the entire population [17]. Costs are either paid by the patient who is reimbursed afterwards (direct payment system, often the case for primary care) or paid directly by the government, except for the copay or nonrefundable part (third-party payer system, mostly for secondary and tertiary hospital care) [17]. Generally speaking, insurance is provided in a hybrid single-payer (with broad coverage) and private insurance system (for additional coverage). Notwithstanding minor differences, the Belgian health care system can be compared with that of France, Germany, the United Kingdom (and most other European countries), and Australia. Although the United States may be different in terms of health insurance, with a stronger emphasis on the private insurance system, it also faces the same questions in terms of apps and wearables.

Currently, there is no fee paid to providers for coaching or follow-up of self-tracking, only for telemonitoring of implanted defibrillators and pacemakers. In 2016, Maggie De Block, the Belgian minister of social affairs and health, announced that she would provide funding for a number of pilot projects in the context of mobile health to investigate the reimbursement for use of specific health apps and devices [18]. In so doing, a matrix needs to be developed for criteria and qualifications, as well as a CE (European conformity) label. For instance, instead of prescribing a sleeping pill, a doctor might prescribe use of a sleep-monitoring app, with the costs subsequently reimbursed by the health insurance.

Results

Overview

Four major themes arise in our body of data:

The patient as health manager: patients are offered numerous possibilities to control and self-manage their health, leading to both opportunities and difficulties. Subthemes are patient autonomy, dropout rates, and the gap between measuring and attaining actual behavioral changes.

Health obsession: the interviewees express concern about a focused use of self-tracking by healthy people, thus creating a “worried well” cohort and widening health disparities. They are critical of the broader medicalization trend in society. Another subtheme is “entertainment medicine,” which refers to questioning the usefulness of digital self-tracking in terms of medical necessity.

Information management: data production, analysis, and interpretation methods change with intensified self-tracking. In this context, providers describe opportunities but also new pitfalls. Subthemes are quality and reliability (of devices and data), importance of context, and data sharing.

Shifts in the roles of the doctor and impact on the health care organization: the impact of digital self-care data on the clinical practice, leading to shifts in terms of data interpretation and the role of the physician. Subthemes are data overload, responsibility, and the importance of in-person contact.

The results show a nuanced and multifaceted understanding of the promises and drawbacks of self-tracking. Findings reveal that digital self-tracking is still emerging and not yet a standard part of the clinical practice, even though the technology is readily available. The interviewed GPs do not often encounter (questions about) new self-tracking technology. In the cardiologist practice, home monitoring (telemonitoring) of defibrillators and pacemakers is already well-established; the cardiologists have firsthand experience with data analysis and complexity of digital health care, although self-care with apps and wearables is not yet an integrated part of their practice and consultation. Just like the GPs in our sample, the cardiologists are not often confronted with patient-initiated forms of self-tracking (eg, digital heart rate monitoring).

Overview of Themes

Theme 1: The Patient as Health Manager

A nuanced sketch of the patient-manager, emerging from the possibilities of self-tracking, is a key theme throughout all the interviews. The digitalization of self-care can be an empowering tool for patients to actively manage their health. Informants agree that classic forms of self-tracking induce a feeling of control in patients that may lead to an increased quality of care.

> *Some people also like to have control over their health and it is not a bad thing that there are methods to meet this demand. Many people are perfectly capable of doing this. You do not need to have studied for nine years to know what high blood pressure or glucose means, and what you can do about it.* [2, GP]

Self-tracking tools allow patients to monitor and adjust their lifestyle personally, which gives them more self-determination and autonomy with regard to their health. Providers acknowledge that self-tracking technologies can offer more insight into bodily information, which opens interesting paths for preventive screening and lifestyle interventions. Patients with unhealthy habits such as smoking can use the technologies as a coach toward a healthier lifestyle. Interviewees see potential in using digital self-tracking to reassure anxious patients, that is, to confirm that there is nothing to worry about.

Still, a majority of our interviewees emphasize that every form of self-care is context-dependent in relation to the specific patient and his or her education and diagnosis. They are careful and nuanced when describing the advantages of the transformation toward a patient-manager and often point to concerns and pitfalls arising with this evolution. We will focus on some of these concerns. Some doctors worry that not every patient is sufficiently skilled to interpret medical data. Another concern focuses on how the “management” role of patients might lead to ignorant self-diagnosis, which already poses problems with classic forms of self-measurements.

> *There is the danger that patients will play doctor themselves. They will themselves decide whether or not to increase their blood pressure medication or diuretic pill.* [1, GP]

Doctor 12 (a cardiologist), however, believes this is not a major problem, as long as patients act within certain limits. For example, patients with diabetes already adjust their medication based on their daily self-tracking of blood sugar levels, which is described as a positive evolution.

Most doctors express concerns about how self-monitoring might lead to more distress and hypochondria by generating complex data and sometimes also information that the patient might not want to be confronted with. This concern reveals a tension with the aforementioned expected reassurance.

> *Another downside is the distress in patients, the problem is that he [the patient] cannot interpret the data himself, so he is alone at home and sees the results, but because of insufficient knowledge, he cannot assess the value of these [results].* [3, GP]

Another significant challenge is the high dropout rate. Our interviewees acknowledge that patients could be burdened with the self-tracking process and quit. Finally, five providers actively questioned the extent to which the apps and wearables actually *improve* health: there is a gap between measuring on the one hand and actual behavioral changes on the other.

> *I have the feeling that they do measure their parameters and that they are subsequently more aware of the problem, but [that] this does not really lead to behavioral changes.* [4, GP]

Theme 2: Health Obsession

When asked which type of patient takes the initiative to self-track, nine out of twelve informants describe how self-tracking is currently mostly initiated by patients who are already healthy: they worry or “obsess” about their health, or they use it to monitor sport activities. Even though GPs and cardiologists are not often confronted with patient-initiated forms of self-tracking, worried patients who do not actually need the tracking technologies for medical reasons most frequently ask questions. Most informants are concerned about this trend.

Although democratization—in terms of availability and easy accessibility of health apps and devices—is perceived as a good thing, most informants are not convinced that patients who are currently hard to reach will suddenly be reached with self-tracking technologies. Instead of fulfilling the promise of democratization, private digital self-care might establish a so-called *Matthew effect* [19]. In economics, this effect refers to the rich getting richer while simultaneously the poor become poorer. In a medical context, this means that the already healthy population might become even “healthier,” whereas the ones who would benefit most from self-monitoring are harder to reach. This raises compelling questions about health disparities.

> *There is an important Matthew effect: those who should not measure, measure, whereas those who do not measure, should measure. Consequently, a lot of money is going to those who do not actually need it, and those who do need it, are not getting it. That is the major problem. The overprotective and already well-controlled patients track themselves and the others do not.* [7, GP]
>
> *Yes, I expect that health disparities might increase because those who will use it [self-tracking tools] are the ones that are already part of the privileged class.* [2, GP]

This also intersects with the concern that the wider dispersal of digital self-tracking technologies might lead to increasing medicalization, that is, framing nonmedical issues, problems, or behaviors in terms of medical problems [20]. Informants are, for instance, hesitant about healthy people who monitor their recreational exercises such as weekly jogging because this monitoring might shift toward medicalization.

Currently, the data of “private” self-tracking are generally not shared with professional medicine, that is, our informants are only rarely confronted with patients who share these data with their care providers. The emerging “parallel circuit” of digital self-care in a home context raises questions about so-called “entertainment medicine” and an overabundance of medically unnecessary data that belong more to the fitness or wellness than to the medical realm.

> *On the one hand I know it [digital self-tracking] will be very useful for certain groups that we currently do not sufficiently reach. But then again, I notice that people who have these technologies now come here to whine...Well, whining might not be the right word...But with these apps you perform a whole lot of ‘entertainment’ medicine.* [7, GP]

Distinctions are subsequently made between self-tracking for fun or motivation versus medical necessity. Having easy access to a visual overview of one’s performances might work as a “motivational tool” (9, cardiologist), but these forms of self-monitoring are generally not medically necessary.

Theme 3: Information Management

This theme focuses on data: self-tracking methods change the way in which data are being generated and interpreted, and this impacts both the patient and the medical practice. In this context, providers discuss the devices and the data, the importance of context, and the extended possibilities for data sharing.

Most interviewees express their struggle with identifying high-quality devices, both in classic and digital material. With regard to classic self-tracking such as a blood pressure monitor, the informants often compare the devices of the patients with their own calibrated devices to look for deviations. Others use lists of approved brands provided by medical organizations. Yet, with regard to digital self-tracking, there are no lists available up to this point. Subsequently, there are high error margins and deviations in quality, reliability, and validity and an overall lack of evidence-based devices.

> *There is also the labeling of these devices. The government must absolutely develop a regulatory framework. This is extremely important, also in terms of technological development. I am very much looking forward to this framework. I notice that a lot of these devices are of poor quality.* [7, GP]
>
> *Certainly, a major problem concerns the validity of these data. That is the basis. With regard to the measurements done with Polars or iWatches, it remains uncertain to what extent these are correct. But these incorrect data are sometimes a reason that patients ask for a consultation. These patients are here because of an incredibly high measurement, but it is unsure whether there is a real problem or just an error.* [10, cardiologist]

Our interviewees are ambivalent about the data overload created by these devices. On the one hand, all informants are critical of this data overload if digital self-care would become an integrated part of professional health care. On the other hand, they acknowledge that more data can lead to more insight into medical conditions. In contrast to a single measurement that takes place at the doctor’s office, digital self-care can result in improved diagnosis, enhanced chronic care, and better preventive care.

Still, another challenge that has an important effect on data interpretation is the lack of context. One example is the fact that devices can potentially contain data of other people. Doctor 12 (a cardiologist) is involved in a project in which heart rhythm is measured with a mobile phone. He once gave the technology to an acquaintance of his who was interested in testing the technology but did not suffer from a heart condition.

> *At one point, a Saturday evening at 11 pm, I received an e-mail that contained a deviated heart rate measurement. I think ‘hmm, this is strange.’ So I send him [the acquaintance] an e-mail and he lets me know that he was at a reception, where he met someone who said that he suffered from a heart rhythm disorder and he [the acquaintance] subsequently gave him his smartphone to try the technology.* [12, cardiologist]

Being unaware of the specific context of the measurement might be potentially dangerous, especially if therapeutic decisions have to be made. This example shows that information always has to be interpreted within a given context.

A final subtheme of information management is data sharing. The diffusion, or sharing, of data, generated from digital self-tracking, can be divided into two categories. First, the self-tracker who shares data on social media, often to inform the network about their progress and to seek motivation. Informants express concerns about privacy infringement and receiving badly informed health advice from members of one’s online network. Second, data can be shared with the doctor and, in turn, with other health care providers through EHRs. If the right balance between useful and unnecessary data can be attained, more centralized data in EHRs could lead to better cooperation and communication with other care providers, a better chronic care, and enhanced quality of care. Problems with privacy, control, and user-friendliness should be tackled first because it requires technological complexity to interpret the data and upload them in the EHRs.

> *It is not just compatibility. It is about privacy. Who has the right to upload data in these files? And who has the right to delete them? It is also about all these questions. It is not that simple.* [8, cardiologist]

Theme 4: Shifts in the Roles of the Doctor and Impact on the Health Care Organization

Digital self-tracking might lead to changes in the roles of the physician in terms of becoming a coach at-a-distance and a data interpreter. Respondents are concerned about the “invasion” of digital self-care data into their practice, questioning the feasibility of the interpretation and usefulness (cf. “entertainment medicine”). Doctor 4 (a GP) draws an analogy between self-tracking and taking a blood test. Instead of only checking those parameters that are required to obtain an answer to a specific medical question, you would just check *all* the parameters the laboratory can possibly examine. As a result, you lose overview and context.

> *If you are going to fill in everything on this form, you will get such a complex picture, with so many results whose outcomes already complicate the interpretation of the question. I believe this situation is similar to self-measurements.* [4, GP]

Some informants express concern about taking responsibility for the data interpretation, given the threat of data overload and a constant flow of data. If patients share the collected data with their doctor, the interpretation of the data and appropriate response becomes the responsibility of the doctor.

> *If I receive all this [these data], I am responsible. It is the same with a blood test: if I have not looked at the data, it is my fault. While if I never received the data, I cannot be held responsible.* [2, GP]

Contrary to self-tracking technologies, no doctor is “always on” to interpret the data, and data interpretation can be time consuming. In *ideal* circumstances, all these data would improve the medical practice, making more time available for an in-depth conversation with the patient. Yet, instead of reducing costs and time investments, the risk exists that data from home monitoring may increase them.

The cardiologists in our sample already encounter problems with data overload in their daily practice. Doctor 8 (a cardiologist), who has experience with home monitoring and implanted defibrillators, observes that they lead to data overload: it is not just registration and monitoring, but the technologies also require data interpretation and responding to it, which is a job that takes 24/7. She states that more cardiologists are required at the hospital to deal with this extra work. Other informants also talk about the need of extra staff such as “telenurses” in a data control room. If paramedical teams can do the first analyses, physicians only become involved in case of actual problems.

When asked, none of the informants were concerned that their authority might be threatened. Most concerns relate to the flow of data and the decrease of “in-person contact” with patients in future medicine.

> *I did not study medicine to sit behind my computer. To have a conversation with the patient will regain importance.* [2, GP]

However, many providers acknowledge that the interaction with the patient will change in the sense that doctors will have to learn to deal with the patient as “patient-manager.” They, among others, point at the more proactive role of patients.

> *The physician must of course be able to deal with this. The generation of physicians, one or two generations ago, started from the model of the dominant doctor: the patient had to listen and the doctor was always right. We already see that this occurs less frequently. Patients ask more questions and discuss more with us, such as discussions related to the treatment.* [11, cardiologist]

This does not mean that our interviewees do not expect benefits from self-tracking for the health care organization. The professionalization and technical automation of self-care might significantly improve professional health care on the condition that problems and challenges are adequately addressed.

> *Especially with chronic patients: that if I make home visits to them, I already have an overview of their self-measurements from the past two weeks, or the last month before I leave. So I can check or look for the best next step, instead of waiting at their home until they have found their written notes or the notes that the nurse wrote down, often in a rush.* [1, GP]

They generally expect the benefits of digital self-tracking to be most obvious in preventive care. Prevention often requires lifestyle changes, and both doctor and apps could help the patient in reaching their individual goals. One cardiologist acknowledges that there is a lot of work to be done on the level of prevention.

> *Regarding acute treatments, not much improvement is required, but in the domain of primary and secondary prevention we can still improve a lot. Secondary prevention also includes lifestyle adjustments, and only the patient himself can attain this. So I think that patient engagement will become very important. And that the doctor becomes more of a coach, instead of a dominant person. And actually, that is the legitimate role of the physician.* [11, cardiologist]

Discussion

Bridging Themes and Literature

The conducted interviews paint an image of new opportunities and challenges instigated by the wider dispersal, accessibility, and affordability of self-tracking tools. Our research findings reveal that self-tracking is expected to lead to shifting understandings of professional care and of the patient–care provider relationship, which is in accordance with other studies’ findings. In what follows, our principal findings are discussed in relation to these studies.

The Healthy, Empowered Patient?

Providers express the expectation that increased self-tracking will lead to more patient empowerment: in their view, self-management of health will gain importance. At the same time, they raise critical questions about the actual results and the possible harms this role can cause to the patient. These two critical points deserve further scrutiny.

From Self-Tracking to Behavioral Change

Five providers describe awareness of a gap between measuring on the one hand and actual behavioral changes on the other. In this context, two questions are currently discussed in the existing literature but remain largely unanswered. First, can wearables affect healthy users’ behavior and promote lifestyle changes? The scientific evidence for this effect is meager [7].

Self-tracking devices are most often used as consumer gadgets, and it remains unclear to what extent they can benefit public health. A recent randomized controlled trial examined the effect of adding wearable technology to a weight loss program [21]. The behavioral intervention turned out to become less effective for 24-month weight loss with the addition of wearable technology. Although there were some limitations to the technology used in this study (eg, the device was worn on the upper arm instead of on the wrist), other studies also struggle with proving long-term effects and face a high dropout ratio. Another recent study examined the feasibility of obtaining measures of cardiovascular health—such as physical activity, fitness, and sleep—via a mobile phone app. Although the observation period in this large-scale study was limited to 7 days, only 9.30% (4552/48,968) of participants who consented actually completed that period [22]. These studies confirm the doubts that our providers express of how to bridge the gap between measuring and an actual sustained behavioral change. Moreover, self-tracking takes place at the individual level, whereas most lifestyle-related problems come with a strong societal connection. Obesity, for example, keeps on increasing for many reasons including an abundance of cheap and unhealthy food and widespread marketing campaigns of the food industry [23]. Deep-seated problems such as obesity cannot be easily fixed by means of self-tracking because they require a structural societal approach.

A second question is whether the use of consumer health apps and devices can improve the outcomes of patients' self-management of chronic diseases. Up to this point, the many expectations about the quantified self for health care are not yet fulfilled [24]. A review study focusing on self-management interventions of rheumatic diseases describes the potential of mobile phone apps but also several pitfalls: scientific evidence for the apps was often lacking, their use can be limited by education level, and the continuous utilization can cause many problems. It concludes that—although online stores offer several apps—more scientific research is needed on the development of such apps [25]. Another meta-analysis focusing on diabetes "identified significant, yet small, reductions in the HbA1c, blood pressure, total cholesterol, and triglycerides levels of patients who were involved in the technology-integrated disease self-management groups" [26]. However, some other patient outcomes did not improve, and it remained unclear why effectiveness was not observed [26]. Although patient empowerment is often read as a precursor to better care and better health, both our interview study and the existing literature point toward a more nuanced view and poses questions about the sheer certainty with which this is accepted.

Patient Harm

This leads to the second part of our discussion of the interviewees' critique: could the role of health-manager cause damage to the patient? Again, we want to focus separately on the healthy individual and the chronic patient. The majority of our interviewees express concern about a focused use of self-tracking by healthy people, creating a "worried well" cohort and widening health disparities. Our finding that mainly already healthy individuals are interested in self-tracking is in accordance with other findings, which shows that mostly young people use smart wearable devices [27]. However, some of the providers see an opportunity here: a small group of early adopters could pave the way toward broader access in the population. The literature, as well as our providers' impression, confirms that digital disparities exist in the adoption and utilization of digital self-care [28]. The aforementioned study focusing on cardiovascular health provides an interesting example, as young male individuals were heavily overrepresented in their study population: of the almost 49,000 individuals consenting to participate, 82.2% were male with a median age of 36 years [28]. Although health IT applications have the potential to address existing health disparities, this requires surmounting several significant technical, practical, and human challenges [28].

With regard to the aforementioned shift toward the patient-manager, informants are critical of the broader societal trend to manage and medicalize health. This excessive emphasis on health is rated undesirable [29,30]. In the literature, it has been described as the "medicalization of health and life itself" [30]. Health is no longer labeled as the absence of disease but has itself become medicalized. This particular understanding of health is linked to the emerging concept of personalized medicine. Many of the digital self-tracking devices are developed within the logic of personalized medicine. The continuous monitoring of healthy people can lead to data overload, overdiagnosis, and overtreatment. The labeling of various aspects of life as "medical" may lead to displacing other values, and the state of being healthy may become impossible to reach [30].

Our providers recognize the increased importance of self-management of chronic diseases while acknowledging the risks that accompany the role of health-manager for patients. For example, nine interviewees explicitly raise concern about increasing distress in patients through self-tracking. They wonder whether patients possess the necessary skill set to deal with the data. Another interview study focusing on patients with multiple chronic conditions raised similar concerns [31]. Two of their findings are of particular interest to our study. First, personal data tracking could carry strong emotional and moral implications: data provoked personal judgments ("good" or "bad") and sometimes negative feelings ("depressed" and "scared"). Second, many patients described self-tracking as a time-consuming effort, even work. In this context, the problem of dropout was also addressed in our study. This concern led some of our interviewees to turn to more paternalistic solutions: to prevent harm, they proposed to limit the patient's autonomy in self-tracking and therefore, to limit the burden of decision making. We question whether this reflex will provide the right answer to the presented challenges; perhaps the solution would be to move forward toward a reviewed concept of patient autonomy. This, however, is beyond the scope of this paper.

Impact on Patient-Provider Relationship

Studies on digital self-care, conducted in the context of experimental studies in Flanders (Veys et al, IBBT TranseCare: deliverable 4.6, 2009, unpublished material) [32], show that patients still value the doctor's expertise but seek more participation. Patients involved in the studies strived for a balanced and reciprocal relationship. They wanted to have a consultation and dialogue with an expert who interprets the data. Although self-tracking could lead to more patient participation, a risk that should be anticipated is that increased patient participation could result in shifting the burden of responsibilities and agency onto patients [33,34]. The focus on self-management, facilitated by digital technology, could damage the clinical relationship by emphasizing the importance of biomedical outcomes in clinical interactions [34].

Overarching Concepts

Finally, we want to discuss two important notions that arose in all our interviews. Although sometimes retreated in the background of the providers' narratives, these notions form an essential part of the nuances that came to the fore in our study. The concepts of context-dependency of health care and the professionalization of self-care are essential to develop a thorough understanding of the providers' report.

No One-Size-Fits-All

The promises and expectations of digital self-care are often grounded in a model of ideal situations. Among other things, they focus on patients who have the right knowledge and adequate skills to interpret the data and to deal with the challenges of these new technologies and who also have financial resources and social opportunity to act on the information. A shared concern of our informants is that this alleged democratization of self-care neglects the context-dependency and complexity of everyday medical practices. Specific uses and success rates are always dependent on multiple variables and patient conditions. Additionally, our research findings show that our informants would not advise *every* patient to track. For example, patients who are expected to suffer from more distress or hypochondria when confronted with all these data would be discouraged to self-track. To facilitate the positive development of digital self-care, academic literature should move away from focusing on "ideal" situations and reflect on more context-relativity and complexity. Both patients and health care providers need guidance regarding these shifts that the digitalization of self-tracking and self-care bring about.

Professionalization of Self-Care

Today, patients are provided with more technologically advanced tools to manage their health. In line with other studies [35], our study illustrates that the boundary between self-care and professional care is decreasing. Kielmann et al [35] focus on the viewpoint of patients, a perspective that our study's focus complements, and show that patients can feel abandoned by the care providers. The patient-care provider relationship must always be a two-way process in which the active role of the patient is acknowledged. Our findings reveal that our informants are critical of not only the quality of these apps and devices but also the reliability of the measurements that the patients themselves have to do. Physicians should not retreat from this evolution toward patient self-tracking but take extra training to deal with data interpretation and to inform their patients. Not all patients are aware of the added value that these technological developments can offer them (eg, sleep monitoring app vs sleeping pills). A good practice of digital self-care requires patient education and digital literacy.

Limitations of This Study

A number of shortcomings of this research must be acknowledged. First, this study addressed a broad theme by incorporating both classic and digital forms of self-care to make their comparison possible. Second, our small sample, although necessary to develop a rich, detailed description, limits our ability to make wider claims based on these interviews. This study explores relatively new phenomena in-depth; further research is necessary to consolidate our findings (cf. infra). Third, we only studied physicians and did not incorporate the viewpoints of patients. Overall, in our small-scale study, we were particularly attentive to implement a number of best practices such as the use of multiple coders, attention to deviant cases, member validation, and a constant comparative technique (cf. supra).

Concluding Remarks and Future Research Endeavors

This study sought to investigate the shifting understandings in medical practice that "private" self-tracking is expected to engender. In so doing, we studied the potential and challenges of self-monitoring while focusing on the everyday professional experience of the physician. Our study offered an analysis of how a small sample of Belgian GPs and cardiologists evaluate self-tracking methods to explore the anticipated shifts. Overall, our research findings showed a nuanced understanding that is in accordance with existing literature on digital self-care.

Our findings open interesting opportunities for other studies. First, our research results can be used for further conceptual scrutiny, especially concerning the presented theme of the patient-manager and the meaning of autonomy in professional health care. As our informants emphasized, the burden of responsibility could prove harmful to some patients. A second interesting path for future research is to explore patients' experiences with digital self-tracking. A main question stemming from our study is whether and how they integrate the data in the patient-doctor relationship. Such studies need to be attentive to context and actual lived experiences. Third, it would be interesting to translate our discussed themes and subthemes to a quantitative questionnaire to reach more doctors and to examine to what extent they share the same ideas. Finally, after analyzing empirical data from both provider and patient perspectives, ethical frameworks should be developed to adequately address the challenges of medicalization of health and professionalization of self-care.

Acknowledgments

This research has been funded by the Research Foundation-Flanders (FWO), grant number G048214N. The authors would like to thank Leah McClimans and Craig Klugman for reviewing a preliminary draft of this manuscript at the ASBH 2016 annual meeting in Washington DC. They also want to express their gratitude to Lisa Parker, Dan Siconolfi, Mark Paterson, and Joel Parthemore for their very useful input and comments. The authors are very grateful to the three reviewers for their valuable comments and suggestions. Finally, they want to thank the participating doctors for the time and effort they spent on this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Informed Consent form.

[PDF File (Adobe PDF File), 23KB - jmir_v20i1e10_app1.pdf]

Multimedia Appendix 2

Interview guide.

[PDF File (Adobe PDF File), 50KB - jmir_v20i1e10_app2.pdf]

Multimedia Appendix 3

Coding Scheme.

[PDF File (Adobe PDF File), 33KB - jmir_v20i1e10_app3.pdf]

References

1. Lupton D. Quantifying the body: monitoring and measuring health in the age of mHealth technologies. Crit Public Health 2013 May 03;23(4):393-403. [doi: 10.1080/09581596.2013.794931]
2. Sharon T, Zandbergen D. From data fetishism to quantifying selves: self-tracking practices and the other values of data. New Media Soc 2016 Mar 09;19(11):1695-1709. [doi: 10.1177/1461444816636090]
3. Godfrey CM, Harrison MB, Lysaght R, Lamb M, Graham ID, Oakley P. Care of self - care by other - care of other: the meaning of self-care from research, practice, policy and industry perspectives. Int J Evid Based Healthc 2011 Mar;9(1):3-24. [doi: 10.1111/j.1744-1609.2010.00196.x] [Medline: 21332659]
4. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. Patient Educ Couns 2002;48(2):177-187. [Medline: 12401421]
5. Topol E. The Patient Will See You Now. New York: Basic Books; 2015.
6. Ossebaard HC, de Bruijn AC, van Gemert-Pijnen JE, Geertsma ME. rivm.nl. The Netherlands: National Institute for Public Health and the Environment; 2012. Risks Related to the Use of eHealth Technologies. An Exploratory Study URL: http://www.rivm.nl/Documenten_en_publicaties/Wetenschappelijk/Rapporten/2013/april/Risks_related_to_the_use_of_eHealth_technologies_An_exploratory_study [accessed 2017-12-07] [WebCite Cache ID 6vUV8lRKo]
7. Piwek L, Ellis DA, Andrews S, Joinson A. The rise of consumer health wearables: promises and barriers. PLoS Med 2016 Feb;13(2):e1001953 [FREE Full text] [doi: 10.1371/journal.pmed.1001953] [Medline: 26836780]
8. Lupton D. Self-tracking modes: reflexive self-monitoring and data practices. 2014 Presented at: ‘Imminent Citizenships: Personhood and Identity Politics in the Informatic Age’ workshop; August 27, 2014; Australian National University, Canberra. [doi: 10.2139/ssrn.2483549]
9. O'Reilly K. Ethnographic Methods. London: Routledge; 2015.
10. Mason M. Qualitative Research. 2010. Sample Size and Saturation in PhD Studies Using Qualitative Interviews URL: http://www.qualitative-research.net/index.php/fqs/article/view/1428/3027 [accessed 2017-12-05] [WebCite Cache ID 6vUX4PUNa]
11. Strauss AL, Corbin J. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Thousand Oaks, CA: Sage; 2008.
12. Boellstorff T, Nardi B, Pearce C, Taylor TL. Ethnography and virtual worlds: A handbook of method. Princeton: Princeton University Press; 2012.
13. Miles MB, Huberman AM. Qualitative Data Analysis: An Expanded Sourcebook. Thousand Oaks, CA: Sage; 1994.
14. Seale C. Validity, Reliability and the Quality of Research. In: Seale C, editor. Researching Society and Culture. London: Sage; 2004:71-83.

15. Creswell JW, Miller DL. Determining Validity in Qualitative Inquiry. Theory Pract 2000 Aug;39(3):124-130. [doi: 10.1207/s15430421tip3903_2]
16. Belgium. Social Security in Belgium URL: http://www.belgium.be/en/family/social_security_in_belgium [accessed 2017-12-05] [WebCite Cache ID 6udBFU04o]
17. Gerkens S, Merkur S. Belgium: Health System Review. Health Syst Transit 2010;12(5):1-266 [FREE Full text]
18. Deblock. 2016. Minister van Sociale Zaken en Volksgezondheid URL: http://www.deblock.belgium.be/nl/maggie-de-block-lanceert-oproep-voor-projecten-rond-mobile-health [accessed 2017-12-07] [WebCite Cache ID 6udBSIHaZ]
19. Joseph KS. The Matthew effect in health development. Br Med J 1989;298(6686):1497-1498 [FREE Full text] [Medline: 2503089]
20. Parens E. On good and bad forms of medicalization. Bioethics 2013 Jan;27(1):28-35. [doi: 10.1111/j.1467-8519.2011.01885.x] [Medline: 21535062]
21. Jakicic JM, Davis KK, Rogers RJ, King WC, Marcus MD, Helsel D, et al. Effect of wearable technology combined with a lifestyle intervention on long-term weight loss: the IDEA randomized clinical trial. J Am Med Assoc 2016 Sep 20;316(11):1161-1171. [doi: 10.1001/jama.2016.12858] [Medline: 27654602]
22. McConnell MV, Shcherbina A, Pavlovic A, Homburger JR, Goldfeder RL, Waggot D, et al. Feasibility of obtaining measures of lifestyle from a smartphone app: the MyHeart counts cardiovascular health study. JAMA Cardiol 2017 Jan 01;2(1):67-76. [doi: 10.1001/jamacardio.2016.4395] [Medline: 27973671]
23. Swinburn BA, Sacks G, Hall KD, McPherson K, Finegood DT, Moodie ML, et al. The global obesity pandemic: shaped by global drivers and local environments. Lancet 2011 Aug 27;378(9793):804-814. [doi: 10.1016/S0140-6736(11)60813-1] [Medline: 21872749]
24. Tomlinson M, Rotheram-Borus MJ, Swartz L, Tsai AC. Scaling up mHealth: where is the evidence? PLoS Med 2013;10(2):e1001382 [FREE Full text] [doi: 10.1371/journal.pmed.1001382] [Medline: 23424286]
25. Azevedo AR, de Sousa HM, Monteiro JA, Lima AR. Future perspectives of smartphone applications for rheumatic diseases self-management. Rheumatol Int 2015 Mar;35(3):419-431. [doi: 10.1007/s00296-014-3117-9] [Medline: 25168866]
26. Or CK, Tao D. Does the use of consumer health information technology improve outcomes in the patient self-management of diabetes? A meta-analysis and narrative review of randomized controlled trials. Int J Med Inform 2014 May;83(5):320-329. [doi: 10.1016/j.ijmedinf.2014.01.009] [Medline: 24534118]
27. Vanhaelewyn B, De Marez L. IMEC. 2016. Digimeter 2016:measuring digital media trends in Flanders URL: http://www.imec-int.com/assets/imec-digimeter-2016-report.pdf [accessed 2017-12-08] [WebCite Cache ID 6vYsl9QwC]
28. Christopher GM. Use of health information technology among racial and ethnic underserved communities. Perspect Health Inf Manag 2011;8:1f [FREE Full text] [Medline: 21307989]
29. Mittelstadt B, Fairweather B, McBride N, Shaw M. Ethical issues of personal health monitoring: a literature review. : Ethicomp Conference proceedings; 2011 Presented at: Ethicomp; 2011; Sheffield, United Kingdom.
30. Vogt H, Hofmann B, Getz L. The new holism: P4 systems medicine and the medicalization of health and life itself. Med Health Care Philos 2016 Jun;19(2):307-323 [FREE Full text] [doi: 10.1007/s11019-016-9683-8] [Medline: 26821201]
31. Ancker JS, Witteman HO, Hafeez B, Provencher T, Van de Graaf M, Wei E. “You get reminded you're a sick person”: personal data tracking and patients with multiple chronic conditions. J Med Internet Res 2015;17(8):e202 [FREE Full text] [doi: 10.2196/jmir.4209] [Medline: 26290186]
32. De Backere F, Ongenae F, Vannieuwenborg F, Van Ooteghem J, Duysburgh P, Jansen A, et al. The OCareCloudS project: toward organizing care through trusted cloud services. Inform Health Soc Care 2016;41(2):159-176. [doi: 10.3109/17538157.2014.965306] [Medline: 25325572]
33. Dedding C, van Doorn R, Winkler L, Reis R. How will e-health affect patient participation in the clinic? A review of e-health studies and the current evidence for changes in the relationship between medical professionals and patients. Soc Sci Med 2011 Jan;72(1):49-53. [doi: 10.1016/j.socscimed.2010.10.017] [Medline: 21129832]
34. Thille P, Ward N, Russell G. Self-management support in primary care: enactments, disruptions, and conversational consequences. Soc Sci Med 2014 May;108:97-105. [doi: 10.1016/j.socscimed.2014.02.041] [Medline: 24632054]
35. Kielmann T, Huby G, Powell A, Sheikh A, Price D, Williams S, et al. From support to boundary: a qualitative study of the border between self-care and professional care. Patient Educ Couns 2010 Apr;79(1):55-61. [doi: 10.1016/j.pec.2009.07.015] [Medline: 19709844]

Abbreviations

EHR: electronic health record
GP: general practitioner
mHealth: mobile health

Edited by G Eysenbach; submitted 12.05.17; peer-reviewed by B Mittelstadt, R Furberg, K Ng; comments to author 09.08.17; revised version received 04.10.17; accepted 04.11.17; published 12.01.18.

Please cite as:
Gabriels K, Moerenhout T
Exploring Entertainment Medicine and Professionalization of Self-Care: Interview Study Among Doctors on the Potential Effects of Digital Self-Tracking
J Med Internet Res 2018;20(1):e10
URL: http://www.jmir.org/2018/1/e10/
doi:10.2196/jmir.8040
PMID:29330140

Original Paper

A Natural Language Processing System That Links Medical Terms in Electronic Health Record Notes to Lay Definitions: System Development Using Physician Reviews

Jinying Chen[1], PhD; Emily Druhl[2], MPH; Balaji Polepalli Ramesh[3], PhD; Thomas K Houston[1,2], MD, MPH; Cynthia A Brandt[4,5], MD, MPH; Donna M Zulman[6,7], MD, MS; Varsha G Vimalananda[2,8], MD, MPH; Samir Malkani[9], MD; Hong Yu[1,2], PhD

[1]Department of Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA, United States
[2]Bedford Veterans Affairs Medical Center, Center for Healthcare Organization and Implementation Research, Bedford, MA, United States
[3]Optum, Boston, MA, United States
[4]Veterans Affairs Connecticut Health Care System, West Haven, CT, United States
[5]Center for Medical Informatics, Yale University, New Haven, CT, United States
[6]Division of Primary Care and Population Health, Stanford University School of Medicine, Stanford, CA, United States
[7]Veterans Affairs Palo Alto Health Care System, Menlo Park, CA, United States
[8]School of Medicine, Boston University, Boston, MA, United States
[9]Diabetes Center of Excellence, University of Massachusetts Medical School, Worcester, MA, United States

Corresponding Author:
Jinying Chen, PhD
Department of Quantitative Health Sciences
University of Massachusetts Medical School
368 Plantation Street
Worcester, MA, 01605
United States
Phone: 1 774 455 3527
Fax: 1 508 856 8993
Email: jinying.chen@umassmed.edu

Abstract

Background: Many health care systems now allow patients to access their electronic health record (EHR) notes online through patient portals. Medical jargon in EHR notes can confuse patients, which may interfere with potential benefits of patient access to EHR notes.

Objective: The aim of this study was to develop and evaluate the usability and content quality of NoteAid, a Web-based natural language processing system that links medical terms in EHR notes to lay definitions, that is, definitions easily understood by lay people.

Methods: NoteAid incorporates two core components: CoDeMed, a lexical resource of lay definitions for medical terms, and MedLink, a computational unit that links medical terms to lay definitions. We developed innovative computational methods, including an adapted distant supervision algorithm to prioritize medical terms important for EHR comprehension to facilitate the effort of building CoDeMed. Ten physician domain experts evaluated the user interface and content quality of NoteAid. The evaluation protocol included a cognitive walkthrough session and a postsession questionnaire. Physician feedback sessions were audio-recorded. We used standard content analysis methods to analyze qualitative data from these sessions.

Results: Physician feedback was mixed. Positive feedback on NoteAid included (1) Easy to use, (2) Good visual display, (3) Satisfactory system speed, and (4) Adequate lay definitions. Opportunities for improvement arising from evaluation sessions and feedback included (1) improving the display of definitions for partially matched terms, (2) including more medical terms in CoDeMed, (3) improving the handling of terms whose definitions vary depending on different contexts, and (4) standardizing the scope of definitions for medicines. On the basis of these results, we have improved NoteAid's user interface and a number of definitions, and added 4502 more definitions in CoDeMed.

Conclusions: Physician evaluation yielded useful feedback for content validation and refinement of this innovative tool that has the potential to improve patient EHR comprehension and experience using patient portals. Future ongoing work will develop algorithms to handle ambiguous medical terms and test and evaluate NoteAid with patients.

(J Med Internet Res 2018;20(1):e26) doi:10.2196/jmir.8669

KEYWORDS

electronic health records; natural language processing; consumer health informatics; usability testing; computer software

Introduction

Background and Significance

Enhancing patient access to their clinical data is a central component of patient-centered care [1,2]. In a nationwide effort to reach this goal [3,4], online patient portals have been widely adopted in the United States to allow patients to interact with their personal health care information, including medication lists and laboratory test results from electronic health records (EHRs) [5]. Initiatives such as *OpenNotes* [6] and the Veterans Health Administration’s (VHA’s) Blue Button [7] also allow patients to access their full EHR notes through patient portals, with early evidence showing improved medical comprehension, health care management, and outcomes [8-11].

However, EHR notes are written for documentation and communication between health care providers [12] and contain abundant medical jargon that can confuse patients [13-18]. In addition, an estimated 36% of adult Americans have limited health literacy [19]. Limited health literacy has been identified as one major barrier to patients’ effective use of their EHRs [5,20-22]. Misinterpretation of EHR content may result in patient confusion about their medical conditions and treatment [23], which could potentially impact service utilization, patient satisfaction, or patients’ own self-management [24].

There has been long-standing research in promoting health literacy [25], including the development of online health education resources, for example. However, these methods do not target clinical notes in an EHR. In addition, health information available on the Internet, although abundant, comes from different resources with varied quality and credibility, which poses great challenges to patients in information seeking and selection [26,27]. The readability levels of health information on the Internet are also often greater than that easily understood by average patients [26,28,29].

A few studies of natural language processing (NLP) systems that translate medical terms to lay terms [30,31] or link them to definitions in controlled vocabularies [32] do show improved patient comprehension [30-32]. These NLP-enabled systems have the merits that they provide patients direct help for EHR comprehension by bundling related health information with individual EHR notes. Despite promising results, these methods have some limitations. First, many medical jargon terms do not have associated lay terms (eg, *neurocytoma* and *lymphangiomatosis*). Second, the definitions of medical terms in controlled vocabularies often contain complex concepts that are not self-explanatory. For example, the medical term *GI* is defined in the controlled vocabulary of National Cancer Institute as “A subject domain utilized for the submission of information encompassing and representing data, vocabulary or records related to gastrointestinal system findings,” where the concept *gastrointestinal* may not be familiar to average patients.

Objective

To address these limitations, we are developing NoteAid, an NLP system that links medical terms in EHR notes to lay definitions targeted at or below the average adult literacy level to support patient EHR comprehension. NoteAid has the potential to be used by veterans supported by VHA, especially the over 2.6 million registered users of Veterans Affairs’s (VA’s) patient portal myHealtheVet [7]. For example, it could be integrated into myHealtheVet as an online tool to help users to understand their clinical notes. Because the challenge in understanding medical terms is not unique to VA patients, NoteAid is also potentially useful for other patient populations. Using a flexible Web-based framework, it can be easily incorporated into patient portals of different health care systems.

In this study, we introduce the main framework of NoteAid and the innovative computational methods we developed to extend its lexical resource and functionality. In addition, as part of the system development procedure, we developed a new evaluation protocol that combines usability testing and quality assessment of lay definitions incorporated into NoteAid by domain experts.

Methods

Study Overview

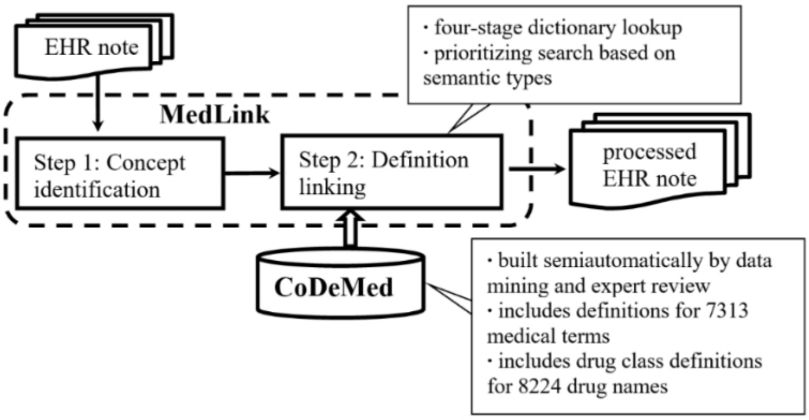
This study presents the NoteAid system and its initial evaluation by physicians. Physicians played the dual role of user and content expert in this usability assessment. Below, we first describe the components and function of NoteAid and then summarize our evaluation protocol.

The NoteAid System

System Overview

NoteAid is a Web application we developed using Java servlets and JavaScript. Figure 1 shows the NLP components and workflow of NoteAid. NoteAid first identifies medical concepts (step 1) and then links them to lay definitions (step 2). NoteAid builds on two core units: CoDeMed, a lexical resource of lay definitions of medical terms and MedLink, a computational unit that links medical terms to lay definitions. We describe CoDeMed and MedLink in the following two sections by focusing on the computational aspects.

Figure 1. Overview of NoteAid. EHR: electronic health record.



CoDeMed

We are developing CoDeMed using both human efforts and automatic methods. By November 2016 when we started this study, CoDeMed contained lay definitions for 7313 medical terms. For example, *Activase* and *bacteremia* were defined as "A drug used to break up blood clots. It is given to patients that have had a stroke or heart attack" and "The presence of bacteria, a type of germ, in the blood," respectively. In addition, CoDeMed included lay language drug class definitions for 8224 medications that did not have term-level lay definitions yet. For example, *Hecoria* and *Neoral* were mapped to the drug class *calcineurin inhibitors*, which was defined as "A drug used to reduce immune response."

For quality assurance, all definitions in CoDeMed were collected from authorized online health education resources (eg, glossaries of National Institute of Health and National Cancer Institute) and simplified and reviewed by domain experts, which included MDs. Because this process is time-consuming, we developed an adapted distant supervision (ADS) system to automatically identify important medical terms from EHR corpora to prioritize the annotation efforts on these terms [33].

We defined important terms as those terms that, if understood by patients, would significantly improve their EHR comprehension. In practice, we used four criteria to judge term importance (details in Multimedia Appendix 1).

Instead of using standard supervised learning, we used distant supervision to save manual annotation efforts. Specifically, we used distant supervision from consumer health vocabulary (CHV) [34] by assuming that medical terms important for patient EHR comprehension must represent medical concepts used by patients. Here, we used the broad sense of "patient" to refer to all the health consumers. CHV contains consumer health terms (which were used by lay people to query online health information) and maps these terms to Unified Medical Language System (UMLS) concepts. As a result, it contains both lay terms and medical terms and links between these two types of terms. Our ADS system used medical terms existing in both CHV and an EHR corpus from the University of Pittsburgh NLP Repository [35] (called EHR-Pittsburgh corpus for convenience) as positive examples and used other candidate terms extracted from this EHR corpus as negative examples to train a classification model. For example, the terms *amyloid*, *hypercholesterolemia*, *laminotomy*, and *pulmonary collapse* were among the positive examples, and the candidate terms *admission blood pressure*, *continued pain*, *fainting*, and *lumbar* were among the negative examples. Training data created in this way had noise. For example, important medical terms that do not exist in CHV (eg, *lumbar*) were labeled wrongly as negative examples. To alleviate this problem, our system used transfer learning and a small amount of manually annotated training examples to adapt the classification model to the target domain to identify medical terms that are important for patient EHR comprehension. We empirically show the effectiveness of ADS by using a gold standard dataset of 6038 EHR terms annotated by domain experts [33]. For each candidate term, the ADS system output its probability of being an important term. We used these probability values to rank candidate terms. The top-ranked terms such as *lipodystrophy*, *myelodysplasia*, and *Parkinsonism* from the EHR-Pittsburgh corpus have been incorporated into CoDeMed.

To improve CoDeMed's coverage, we developed an unsupervised method to mine medical synonyms from Wikipedia [36]. Specifically, we used the interwiki links in pages on the Wikipedia Health tree to extract candidate synonyms for medical terms. For example, *avian influenza* and *bird flu* are linked to the term *avian flu* through the interwiki links, which are both good synonyms of this term; 15 terms are linked to *blood pressure*, which include the synonym (*BP*), the hyponyms (eg, *systemic blood pressure* and *diastolic blood pressure*), and other related terms (eg, *blood pressure measurement* and *low blood pressure*). We then ranked these candidates by using word embedding and pseudo-relevance feedback. Word embeddings are distributed representations of words (which typically are high-dimensional real-valued vectors) learned from large unlabeled text data. Words sharing similar semantics and context are expected to be close in their word vector space [37]. Pseudo-relevance feedback [38] is a method widely used in information retrieval to obtain a better representation of a target concept by using the retrieved results as pseudo-relevance feedback (as opposed to relevance feedback from human annotators). In our case, the target concept is a medical term for which we sought synonyms. We first used other strategies (eg, cosine similarities between the target term and its candidate synonyms) to rank the candidates. We then used the top-ranked candidates to represent the target term and reranked the candidates. We evaluated our methods on 1507 synonyms and nonsynonyms manually judged for 256 medical terms [36]. This method has been used to enrich the candidate synonym set for CoDeMed.

MedLink

We developed MedLink to retrieve lay definitions from CoDeMed for medical terms in EHR notes. MedLink utilizes MetaMap [39] to identify medical terms and implements a linking function to retrieve their definitions from CoDeMed.

MetaMap is a widely used lexical tool developed by National Library of Medicine that automatically maps medical text to medical concepts in the UMLS Metathesaurus [39]. For example, both *Cushing's syndrome* and *Cushing disease* are mapped by MetaMap to the UMLS concept *Cushing Syndrome*.

We have developed two strategies to improve MedLink's definition linking function. The first strategy is a four-stage dictionary lookup procedure that uses synonym finding and partial string matching to improve coverage of medical terms. For each term identified by MetaMap, MedLink first searches CoDeMed for this term by exact match. If the search fails, MedLink searches for the term's UMLS preferred name, which is identified by MetaMap. A search failure leads to the third stage, where MedLink shortens the term by removing words in a list (Multimedia Appendix 2) of common modifiers of diseases and body locations, such as *chronic*, *severe*, and *left* and searches the trimmed term. If the search fails again, MedLink searches individual words in the trimmed term and displays lay definitions for every single word that has a hit in CoDeMed. To improve search efficiency, MedLink uses a hash table to store interim search results for reuse.

The second strategy is to speed up the system by prioritizing or deprioritizing EHR terms by their UMLS semantic types. Specifically, we defined 21 prioritized semantic types and three deprioritized ones (Multimedia Appendix 2) by extending our previous work [32] using feedback from domain experts. Prioritized semantic types represent semantic categories of medical concepts commonly used in the clinical domain, such as *Disease or syndrome*, *Pharmacologic substance*, and *Laboratory or test result*. Deprioritized semantic types represent concepts that are too general to have a standalone clinical meaning, such as *Geographic area* and *Temporal concept*. Terms with deprioritized semantic types are ignored (ie, not translated by the system). Terms with prioritized semantic types are searched in CoDeMed using the aforementioned four-stage dictionary lookup procedure. The remaining terms are searched in CoDeMed at only the first two stages of dictionary lookup.

Evaluation Protocol

Content for Evaluation Protocol: Electronic Health Record Notes

The EHR notes used for this evaluation were chosen from the EHR-Pittsburgh corpus because this corpus was deidentified and available for research [35]. Specifically, we randomly selected 200 progress notes from this corpus that satisfied the following two criteria: (1) containing the Assessment and Plan section and (2) containing at least ten medical terms as identified by MetaMap. An expert in public health, who has worked in the civilian and military health care fields for 20 years in the specialties of dermatology, surgery, and emergency medical services reviewed these notes sequentially and selected the first 10 notes whose Assessment and Plan sections include good narratives. Three criteria were used to identify good narratives: (1) the text was written in the conversational tone, as opposed to bulletin points often used in a review of systems or update for other health care providers; (2) the text contained important information about a patient's diagnoses and treatment plans; and (3) the text was not trivial and contained at least five medical jargon terms. We chose the Assessment and Plan section because this section often contained content that satisfied the first two criteria. We used these sections of the selected notes for system evaluation. The notes we selected were mostly (9 out of 10) intensive care unit (ICU) notes. We chose those notes because they contained abundant complex medical jargon that could be used to challenge the system to test its robustness.

Textbox 1 shows an excerpt from one clinical note used for system evaluation, with a number of medical terms that may hinder patients' comprehension italicized. Here we show a subset of terms identified by the UMLS lexical tool MetaMap [39] for illustration purposes.

Usability Procedures

Our protocol allows physicians to simultaneously assess the system's user interface and the content quality of lay definitions. This approach is motivated by two factors. First, our system provides patients knowledge of medical terms. The quality of the provided knowledge is an important aspect of its usability. Physicians rather than patients have the proper training to judge the accuracy of definitions for medical terms. Therefore, we asked physicians to evaluate our system at this stage. Second, physicians do not differ from lay people as users of computer software. Here, we used the general meaning of "users of computer software," that is, people who use a software product without the technical expertise required to fully understand it. From this perspective, we expect physicians to give feedback on user interface (eg, ease of use and speed) in a similar way as patients do.

The evaluation included a 1-hour cognitive walkthrough session and a 7-item postsession questionnaire. Each physician was interviewed separately in the following procedure. At the beginning, the interviewer gave the physician an overview of the system, the assessment procedure, and the goal of the interview—collecting feedback regarding the system's user interface and output content. In addition, she informed the physician that the system output might not be always accurate and encouraged the physician to seek clarification on definitions they found inaccurate or confusing. She then showed the physician instructions on system use. The physician used the system to process the EHR excerpts one by one, reviewed the output from the system, and gave feedback in a think-aloud manner. The physician was encouraged to make free comments on any aspect of the system.

Textbox 1. Illustration of medical terms in a sample clinical note.

Cardiac—The patient was *hypotensive* yesterday during the day with *pressures* running in the *systolics* of 80's to 90's by *cuff*. *Cardiology* was called to see the patient and they did a quick bedside *echocardiogram* that revealed no *pericardial effusion*. Her *troponins* never went higher than 0.77 and *cardiology* was not concerned with any primary *cardiac event*. Her heart rate was also in the one teens to one twenties.

During this process, the interviewer also asked the physician a few optional questions about the user interface (details in the subsection Physician Responses to Prompts). Except these prompts, the interviewer only observed the physician using the system and responded to his or her questions, without discussing or debating on suggestions for system improvement.

The postsession questionnaire was developed by a group of experts in public health, medicine, health informatics, and computer sciences. Because this study is the first effort to collect physicians' feedback on the content quality of the NoteAid system, there are no existing validated surveys to use. We therefore elected to develop a short survey with questions that were specific to our system. To ensure the quality of the survey, we asked one clinical domain expert to evaluate the validity of the survey content and asked two lay people to evaluate whether the content is easy to understand. Our questionnaire includes 5 scale questions and 2 optional open-ended questions. The scale questions (details in the subsection Physician Responses to Postsession Questionnaire) evaluate lay definitions in four aspects: readability (Q1), informativeness (Q2), coverage (Q3), and accuracy (Q4 and Q5). The open-ended questions collect free comments on any aspect of the system (details in Multimedia Appendix 3).

We recorded voice, screen, and mouse clicks of the whole interview process, including filling out the survey, by using Morae Recorder (Version 3.3.4., TechSmith Corporation) for data analysis.

Participants

A convenience sample of 10 physicians with diverse clinical expertise (details in Table 1) were recruited from Edith Nourse Rogers Memorial Veterans Hospital, VA Palo Alto Health Care System, VA Connecticut Health Care System, and the University of Massachusetts Medical School.

We used a small group of physicians by following previous studies in usability research [40-44] and evaluating clinical NLP systems [30,45-47] and considering factors relevant to our case. The variability of the physicians' specialty was unintended. Previous work found that, for simple usability tasks, usability testing using 5 to 10 users was able to find over 80% problems, and the cost-benefit ratio of increasing the number of users was high [40-43]. Our task was simple and required the user to do a small, closed set of operations (eg, copying and pasting EHR content into the input box and hitting the "simplify" button to view the output). In addition, our study is a midstage evaluation to prepare the system for late-stage patient evaluations.

Data Analysis

We analyzed the audio-recorded think-aloud data and physician responses to open-ended survey questions by qualitative content analysis. Qualitative content analysis is a research method widely used for analyzing written, verbal, or visual communication messages through the systematic process of coding and identifying themes or patterns [48-50]. It has been successfully used to study clinical NLP systems [45,46] and patient's comprehension of clinical text [14,23]. Following established techniques [49,50], we carried on the analysis over three phases, that is, preparation, organizing, and reporting.

In the preparation phase, two researchers reviewed the think-aloud data from two interview sessions and identified physicians' comments related to system's user interface and lay definitions. This review resulted in an initial code book with three broad top-level categories: feedback about the user interface, feedback about the definitions, and other feedback. The first category was further divided into positive comments, system error, and suggested improvements at the second level. The second category was divided into positive comments, inaccurate definitions, suggested improvements, missing definitions, and lay terms that do not need definitions.

In the organizing phase, one researcher coded the data, which were then reviewed by the second researcher. The coder segmented the data, classified the segments by selecting codes (details in Multimedia Appendix 4) from the existing codebook, and created new codes when necessary by discussing with the reviewer. Codes were assigned based on the manifest content of the recordings. Code definitions and coding examples were used to facilitate the coding process (details in Multimedia Appendix 4).

Table 1. Specialty and gender of physicians who participated in NoteAid assessment.

Specialty	Gender	
	Female, n	Male, n
Clinical Pharmacology		1
Endocrinology	1	1
Family Medicine		1
Internal Medicine	2	2
Preventive Medicine	1	
Pulmonology	1	

Few disagreements (4.5% [33/728]) over the coded data were resolved by discussions between the coder and the reviewer. The final coding scheme and categories (ie, themes) formulated over the codes are summarized in Multimedia Appendix 4. To assess intercoder reliability, two researchers independently coded 229 segments from physicians' think-aloud data, which covered two to four notes randomly selected for each physician. The intercoder agreement on this dataset is .88 Cohen kappa.

In the last phase, we reported descriptive statistics of themes and qualitatively summarized key findings.

Results

System Output

Figure 2 shows a snapshot of NoteAid's output on the EHR excerpt in Textbox 1, where 8 medical terms were highlighted and linked to lay definitions by the system. The definition of a medical term, for example, *troponins*, will show up when hovering over this term. The definitions output by NoteAid for all the 8 medical terms are listed in Multimedia Appendix 5.

Table 2 summarizes the EHR data used in the evaluation and NoteAid's output on this data. As shown in Table 2, NoteAid linked 29.2% of the terms identified by MetaMap to lay definitions. The low ratio is caused by two reasons: (1) some medical terms do not yet have lay definitions in CoDeMed and (2) MetaMap identified both medical jargon terms and lay terms (eg, *patient*, *bleeding*, and *elevated*) commonly used in the biomedical domain. Therefore, this ratio value underestimates NoteAid's recall. In our result analysis, we used medical terms that at least two physicians judged to be missed by NoteAid and medical terms linked by NoteAid to estimate the recall, which was 0.565 (standard deviation 0.164) per note in this study.

Analysis of Physician Think-Aloud Data

Overview

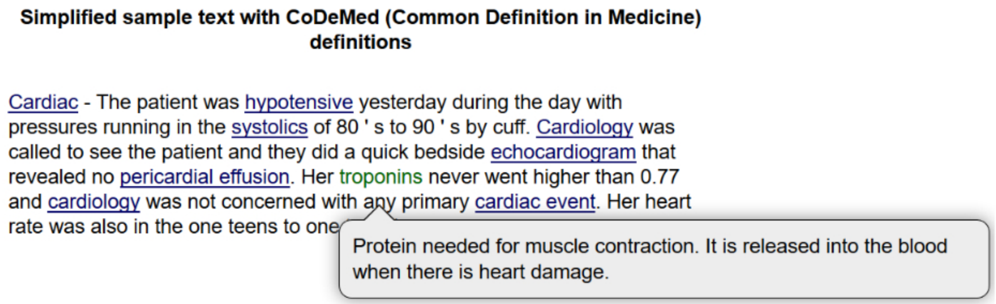
In total, 8 physicians reviewed 10 EHR notes; 2 physicians reviewed 9 notes in their interview sessions and could not stay longer. We asked all the physicians to fill out the postsession questionnaire even if they did not finish all the notes. We used all available think-aloud data for analysis and coded 728 segments including 71 system-related comments, 593 definition-related comments, and 64 other comments.

System Related

The system-related segments include 11 positive comments (1.1), 50 comments suggesting improvements (1.2), and 10 system errors (1.3). Textbox 2 shows some examples. The numbers in the round brackets are the codes assigned for these categories, as detailed in Multimedia Appendix 4. In the rest of this paper, we used A#'s (eg, A1, A2) to represent different physicians.

One major suggestion for improvement (25 comments from 7 physicians [A1, A3, A4, A5, A7, A8, and A9]) was to change the display of definitions for partially matched terms. NoteAid currently highlights the full spans of those terms and displays definitions for their component words or subterms, hoping to give patients as much information as possible. For example, *normal sinus rhythm* is highlighted, but only *rhythm* is defined (which is displayed as "[rhythm]: definition text"). However, many physicians thought this was somewhat confusing. A deep analysis of other comments for the user interface suggests the following improvements: (1) showing system status when it is running (A2, A4, A5, and A6), (2) putting the " simplify" button under the search box (A2, A4, A5, and A6), (3) making the label of the search box more informative (A4 and A6), and (4) disabling the dangling hyperlinks (which we planned to use to link to educational materials in the future) for now (A2 and A6).

Figure 2. A snapshot of system output.



Simplified sample text with CoDeMed (Common Definition in Medicine) definitions

Cardiac - The patient was hypotensive yesterday during the day with pressures running in the systolics of 80 ' s to 90 ' s by cuff. Cardiology was called to see the patient and they did a quick bedside echocardiogram that revealed no pericardial effusion. Her troponins never went higher than 0.77 and cardiology was not concerned with any primary cardiac event. Her heart rate was also in the one teens to one

Protein needed for muscle contraction. It is released into the blood when there is heart damage.

Table 2. Statistics of the electronic health record (EHR) data used for evaluation and NoteAid's output on these data.

Characteristics of NoteAid's output	Data
Number of EHR[a] excerpts, N	10
Number of words per EHR excerpt, mean (SD)	383 (169)
Number of terms identified by MetaMap per EHR excerpt, mean (SD)	141 (66)
Number of terms linked to lay definitions by NoteAid per EHR excerpt, mean (SD)	41 (19)
Percentage of terms identified by MetaMap per EHR note that have been linked to lay definitions by NoteAid, mean (SD)	29.2 (2.8)

[b]EHR: electronic health record.

Textbox 2. Examples of physicians' comments on NoteAid's user interface. A#'s refer to different physicians. The numbers in round brackets, for example, (1.1), refer to the codes assigned for data categories.

Positive comments (1.1):

A1: "It's very efficient in terms of linking the term to the definition. It's fast and accurate as well."

A4: "The system on this first note, the neuro note, is doing a good job."

A6: "This could be a great product for pre-med or early med students."

A7: "You picked up the meds (medications) and the drugs well, most of the concepts."

A10: "Overall I think it's nice how it's done this."

Comments suggesting improvements (1.2):

A2: "Ok, so in that case, maybe when someone hits it once, give a, like, time bar or something."

Interviewer: "To let them know it's working?"

A2: "Yeah." (1.2.2)

A4: "Again, you're underlining more than you actually define." (1.2.1)

A5: "Maybe the simplify button should go right under there." (1.2.2)

A6: "Oops, what happens when I click on it, it just hops away. Why does it do that?" (1.2.2)

A7: "What is this?"

Interviewer: "I think that's something they were trying, I'm not sure if it's going to help or just make people more confused, to have multiple senses."

A7: "Yeah, that was confusing, actually." (1.2.2)

Definition Related

Among the definition-related comments, 28 were statements about good definitions (2.1), 54 were about inaccurate definitions (2.2), 92 were about improvements of definitions (2.3), 18 were about lay terms that should not be defined (2.4), and 401 were about missed terms (2.5). Textbox 3 shows some examples. As one term can occur in multiple notes and be commented on by different physicians, we further analyzed these data at term level and identified terms using agreement by at least two physicians.

We found that physicians usually passed good definitions silently, resulting in fewer positive comments (see three examples in Textbox 3).

A total of 12 terms (ie, 6.2% of unique terms linked to lay definitions by NoteAid) were judged to be inaccurate, which included 11 ambiguous terms and acronyms whose definitions provided by NoteAid did not fit the specific context. For example, *AC* is defined by NoteAid as "a short-hand name for a chemotherapy combination used to treat breast cancer," but it was used as an abbreviation of *assist control* in the EHR note.

One definition (for the term *Valcyte*) was judged to be inaccurate despite the context. This definition was derived from drug class definitions to improve NoteAid's coverage of medical terms. Specifically, NoteAid treats *Valcyte* as one type of *Purine Nucleosides*, which is defined as "a drug used to treat cold sores, genital herpes, and chicken pox."

Physicians suggested improving definitions for 20 (10.3%) terms. In particular, they suggested adding information to the definitions for three terms (ie, *sinus rhythm*, *diuresis*, and *Lantus*) to improve their clarity and specificity. For example, *sinus rhythm* is defined as "Heart rhythm that begins at the upper chamber of the heart" in CoDeMed, whereas physicians wanted us to clarify that this term referred to the normal heart rhythm. Other suggestions include improving the grammar and readability of certain definitions and unifying the style and granularity of the drug definitions.

A total of 91 terms (ie., 14.8% of unique terms in the ten EHR notes that were identified by MetaMap) were judged to need lay definitions but missed by NoteAid, which include 34 multi-word terms (eg, *tidal volume*, *GJ tube*, and *community acquired pneumonia*) and 57 single-word terms (eg, *FiO2*, *macrocytic*, and *reintubated*). The 34 multi-word terms have meanings beyond the simple sum of their component words. For example, knowing the words *community*, *acquired*, and *pneumonia* is not sufficient for understanding the term *community acquired pneumonia*.

Two terms, *felt* and *level*, were linked to lay definitions but were judged by domain experts to be lay terms that did not need definitions.

Other Comments

Comments in this category are mostly observations or suggestions to the evaluation. For example, three physicians (A1, A4, and A5) commented that certain notes were too complicated for patients to understand even if we provided lay definitions for medical jargon, and one physician (A5) suggested the inclusion of simpler outpatient notes for future evaluations. Another noticeable pattern is about note comprehension. For example, two physicians (A3 and A4) pointed out that, in addition to showing the definitions of lab measures, explaining the range of the normal lab values would be also important for patient understanding their lab results.

Textbox 3. Examples of physicians' comments on definitions that NoteAid provided for medical terms. A#'s refer to different physicians. The numbers in the round brackets, for example, (2.1), refer to the codes assigned for data categories. The context for the term wean in the fourth example is "I will continue to wean his FiO2 as tolerated and will attempt a trach-mask trial today." NoteAid did not link the term hemodynamically in the fifth example to a definition.

Term: antibiotics

Definition: Medicines that attack bacteria.

A5: "I like the focus on bacteria there." (2.1)

Term: BUN

Definition: [full name: blood urea nitrogen] The amount of nitrogen in blood is measured to check how well the kidneys work. The nitrogen comes from urea which is formed by the breakdown of protein in the liver.

A9: "I like how this one talks about how it's measured to check how well the kidneys work..." (2.1)

Term: hemoglobin

Definition: A protein in the blood that carries oxygen. It gives blood its red color.

A4: "It talks to people at the level they can understand, and this extra little bit...I feel like that sort of helps you feel like, I get that. It's just well done." (2.1)

Term: wean

Definition: To be taken off a certain drug very slowly.

A2: "Wean here, the definition is to be taken off a certain drug so I would suggest maybe to change it to be taken off something slowly or a certain regimen." (2.2.2)

Term: hemodynamically

A3: "Hemodynamically needs to be defined." (2.5.1)

Term: community acquired pneumonia

Definitions: [[community]]: A group of people. [[pneumonia]]: An infection of the lungs, usually caused by viruses or bacteria.

A4: "Community acquired pneumonia shouldn't be defined as two separate things, there's no point in defining community for people, and CAP is a concept and should be defined together." (2.5.2)

Term: felt

Definition: Feeling happy, mad, or scared.

A5: "I don't think we need to define felt." (2.4)

Term: creatinine

Definition: A waste product made by muscles and cleared from the blood by the kidneys.

A8: "This again should tell the reader clearly that these are tests that are used to monitor how well the kidneys are working." (2.3.1)

Physician Responses to Prompts

Table 3 summarizes physician responses to optional prompts, manually labeled as *satisfied* and *suggesting improvements*. Most respondents thought the system's speed acceptable. One respondent (A9) commented that higher speed would be better for batch processing of many notes at one time. One respondent (A5) suggested showing a progress bar when the system was working. One respondent (A4) suggested disabling the "simplify" button to avoid multiple hitting (which would slow down the system) when the system was processing a note. All respondents liked how NoteAid displays definitions, but one (A4) commented that using hyperlinks could cause confusion. Most respondents thought that the system was easy to use.

Physician Responses to Postsession Questionnaire

Table 4 summarizes physician responses to scale questions. The average scores for the four aspects of lay definitions generated by NoteAid range between 3.70 and 4.30.

The analysis of physician responses to the open-ended questions (details in Multimedia Appendix 3) showed that physicians would like to see improvements in robustness (A3 and A6) and layout (A7, A9, and A10) of the user interface and quality of certain aspects of definitions (A2, A4, A6, and A8). In addition, two physicians (A4 and A5) suggested using outpatient notes for future system testing.

Among the positive comments, five physicians (A1, A3, A5, A8, and A10) thought that the user interface was very straightforward and easy to use. Two physicians (A1 and A2) liked the good coverage of medical terms, and one physician (A9) appraised the lay language nature of definitions. In addition, three physicians (A4, A6, and A9) appreciated the usefulness of the system.

Table 3. Physician responses to optional prompts on user interface. We report the proportion of physicians who were satisfied or suggested improvements when responding to each prompted question. A#'s denote different physicians.

Questions[a]	Satisfied	Suggesting improvement	Example responses
PQ1. Do you feel the system speed is tolerable?	4/7 [A5, A6, A8, and A10]	3/7 [A3, A4, and A9]	A3: "I think it's ok. I mean, some of them, this one was a long one and it's taking a little longer. But I suppose people would probably figure that out." A5: "If there was a progress bar or something it would be okay." A6: "Yeah, it seems really fast today." A9: "I think it takes a little longer than ideal to process it. Well just cause the longer it takes, especially when people are trying to do multiple things. But it's not too bad."
PQ2. Do you like the way that the system displays the definitions?	9/10 [A1, A2, A3, A5, A6, A7, A8, A9, and A10]	1/10 [A4]	A3: "I like it fine. I think it's very easy to see which terms are defined and to see the definition. It's pretty straightforward." A4: "Because there's an underline, I want to click on it. If you had it blue but not underlined...I don't know what to tell you guys to do, it definitely says to me "click on me" and if I'm moving quickly as people are, like, I would figure it would but it's a little confusing." A5: "Yeah, yeah I do. Just hovering over with the mouse. It is convenient." A7: "Yeah, they're short and easy to read."
PQ3. Is it easy for you to find the definitions that were generated by the system?	4/4 [A1, A2, A3, and A7]	0/4	A1: "Definitely, I just have to point the mouse at any word that I'm wondering what it means." A7: "Oh yeah, you just hover, that's not hard."
PQ4. Do you think the instructions on the web page are easy to follow?	7/8 [A1, A2, A3, A5, A7, A8, and A10]	1/8 [A4]	A1: "Yes, very easy" A4: "I didn't follow them, I just did what you told me. So, you haven't labeled the search box, so I would...or 'into the box below.' That may be confusing, I don't know...I would probably give some direction like 'push simplify button.'" A10: "Yeah, just copy and paste."

[a] PQ# refers to the #th prompted question.

Table 4. Evaluation scores for scale questions in postsession questionnaire.

Questions	Domain	Scale[a]	Evaluation score Mean (SD)
Q1. The definitions are in lay language (ie, do not contain medical jargon).	Readability	never 1 2 3 4 5 always	3.90 (0.57)
Q2. The definitions provide useful information for comprehending medical jargon in electronic health record (EHR) notes.	Informativeness	never 1 2 3 4 5 always	3.80 (0.63)
Q3. NoteAid has good coverage of lay definitions for medical jargon in EHR notes.	Coverage	disagree 1 2 3 4 5 agree	4.10 (0.74)
Q4. NoteAid links medical jargon to definitions that are correct or appropriate for patients.	Accuracy	seldom 1 2 3 4 5 often	3.70 (0.67)
Q5. NoteAid links medical jargon to incorrect definitions.	Accuracy	often 1 2 3 4 5 seldom	4.30 (0.95)

[a] We used 5-point Likert-style scale questions. For example, "never 1 2 3 4 5 always" refers to never, seldom, sometimes, often, always.

Discussion

Principal Findings

We have developed NoteAid, an NLP system that links medical jargon to lay definitions targeted at or below the average adult literacy level and reported a formative evaluation conducted to improve the system. Ten physicians with diverse backgrounds evaluated NoteAid. Overall, physicians were positive about the user interface and the quality of lay definitions, as indicated by their responses to the survey questions and optional prompts.

Building NLP systems that support patient EHR comprehension is challenging. One major reason is the language use characteristics of EHR notes related to clinicians' writing behavior [51]. For example, clinicians often use shorthand (eg, abbreviations and acronyms) in clinical narratives, which causes great ambiguity in meanings [52,53]. Simple strategies such as the most frequent sense (MFS) heuristic, is not sufficient to resolve such ambiguity (details in the subsection Lay Definitions). Misspellings are common in clinical texts, which reduce the system's recall on medical terms. Some clinical texts are ungrammatical, for which lexical-level comprehension support is not sufficient. These characteristics also make it harder for patients to comprehend their EHR notes, given that the notes are already abundant with medical terms. Despite these challenges, NoteAid received positive physician feedback, suggesting that this tool has great potential to be useful for patients.

In addition, physicians provided valuable suggestions for system improvement before we started a randomized trial sponsored by VHA, which fulfilled one major goal of this study.

User Interface

Physicians thought that system speed was satisfactory (although not optimal) and that the system website was easy to navigate. The definitions, which pop up in grey bubbles when the cursor is hovered above a term, were helpful and easily found. Simple changes such as altering the placement of buttons on the website or revising the label of the input box to add clarity were suggested. We have improved the system based on their feedback in five aspects.

First, we fixed the position of the “simplify” button to be right below the input box, so a user can easily find it even when the output text is long.

Second, we added a function to disable and grey out the “simplify” button when the system starts to process the input text and reenable it after the processing completes. This gives users a sense about when the system is working. It also prevents them from hitting the “simplify” button repeatedly, a behavior that can slow down the system.

Third, we modified the introduction text on the user interface and the label of the input box to improve clarity.

Fourth, we disabled dangling hyperlinks in system output.

Fifth, we investigated the system errors that occurred during the evaluation and found that the errors were caused by a design flaw of the original system. Previously, the system segmented an input text into sentences and reopened a subservice of MetaMap every time it was processing a new sentence. Consequently, the temporary files opened for communication with the subservice sometimes accumulated too quickly to be tolerated by the operating system, causing a system crash. We have fixed this problem by modifying the code to process the input text without sentence segmentation. This change also improved the system speed.

We provide screenshots of the original system and the enhanced one in Multimedia Appendix 6.

Our system has the potential to be integrated into VA’s patient portal after patient testing and further improvement. It can also be used as a standalone tool. Using our system only requires a few simple operations; therefore, we expect it to be easily used by patients with different levels of computer skills. We implemented the system using Hypertext Transfer Protocol Secure that encrypts user page requests and responses from the Web server, to protect patients’ private information.

Lay Definitions

Physicians often passed good definitions silently. Their positive comments suggest that they liked our approach for defining medical jargon—giving essential information in lay language while omitting other details that could easily overwhelm the patients. This approach helps reduce patients’ cognitive load when they read long EHR notes abundant with medical terms.

Furthermore, physicians gave valuable feedback about which medical concepts are better understood as phrases versus the sum of their individual terms. We have used their feedback to improve our guideline on selecting compound medical terms for inclusion into CoDeMed. This improvement will reduce cases where patients think they understand a term like *community-acquired pneumonia* but in fact not, therefore helping them understand their EHR notes more accurately. The compound terms physicians identified can also be used as examples to develop learning-based methods for automated compound term detection, which will be our future work.

They also identified several inaccurate definitions that need immediate attention. The inaccuracies were mostly caused by ambiguity. We identified two major sources of ambiguity by result analysis: (1) drugs that can treat different diseases and (2) acronyms. NoteAid currently uses the MFS heuristic for acronyms, which were judged by domain experts, and outputs all definitions for multi-purpose drugs. However, the evaluation results suggest that MFS is not accurate enough, and showing multiple definitions for a drug name could cause confusion. Improving accuracy is crucial for patient use of NoteAid because inaccurate definitions add confusion to patients, and it would be unlikely for them, especially those patients with low health literacy, to discover such inaccuracies by themselves. In the future, we will address this problem by developing automatic methods to predict senses of ambiguous terms in EHR notes.

From physician’s comments, we observed some disagreement about appropriate length and scope of definitions. Most physicians agreed that longer, more detailed descriptions of medical concepts were preferable; however, one physician (A6) argued for the *less is more* approach, citing the already too complex nature of EHR notes. In addition, some physicians thought that only providing lay definitions to medical terms may not be sufficient for full EHR note comprehension. These comments suggest that patient EHR comprehension is challenging and unlikely to have a one-shot solution. For example, patients at higher literacy levels may prefer more detailed definitions, whereas patients at low levels may find such definitions overwhelming. This feedback suggests that the current NoteAid design may be sufficient to support EHR comprehension, but in some cases additional tools may be required. For a more comprehensive solution, one may extend NoteAid by incorporating definitions at different granularity levels and methods interactively measuring patients’ literacy levels.

Drug definitions were another source of disagreement between physician reviewers. Although some preferred simple, all-encompassing definitions (eg, all antibiotics having the same brief definition, “ *Flagyl: a drug that kills germs* ”), others felt more detail should be provided about specific conditions each drug is used to treat. One way to resolve the discrepancy is to introduce hierarchical definitions that show the simple generic definition first and then details specific to context.

Physicians identified many terms which do not present in CoDeMed and thus were missed by NoteAid. We have added 4502 more definitions in CoDeMed since this evaluation, which improved NoteAid's recall of medical terms in the ten EHR excerpts used for this study from 0.565 to 0.787. We have also developed a hybrid strategy that uses the UMLS controlled vocabularies to back up CoDeMed for missed terms while we simultaneously continue to improve its coverage. We expect that including more lay definitions will benefit patients, especially those with low health literacy.

Related Works

In the Introduction section, we have discussed the difference of our system from previous systems [30-32] in supporting patient EHR comprehension. Our system is the first one that links medical terms to lay definitions in EHR notes. Our system used drug class definitions for medications that did not have term-level definitions in CoDeMed, which shared the same spirit with the method of using UMLS relations (eg, "Is-A" relation) to generate general explanations for medical terms that lack corresponding lay terms [30].

Previous work in evaluating NLP systems supporting patient EHR comprehension focused on their effects on comprehension [30-32] and sometimes cleaned system outputs manually before presenting them to users [30]. The purpose of our evaluation is different. We evaluated both the user interface and content quality of the system's original output as part of our development effort to prepare the system for a late-stage patient evaluation.

Similar to previous work in evaluating clinical NLP systems [30,45-47], we used task-based evaluation. The think-aloud walkthrough evaluation protocol used in our study is similar to previous work in evaluating online search engines in answering physicians' questions [45] and clinical NLP software [46]. Zheng et al [46] found that clinical NLP software that required installation before use often presented challenges to end users. Compatible with their findings, our system, which adopts a Web-based framework without installation requirement on the user side, received positive feedback in terms of ease of use.

Limitations

Similar to previous work in usability testing of clinical NLP systems, we obtained feedback about NoteAid from a small group of users. Usability testing identified important issues for system improvement, although a larger scale evaluation may yield additional refinements. We used notes from ICU patients, which may not be representative of notes for patients at lower acuity, including individuals seen in outpatient clinics. However, because these notes often cover complex clinical issues, they offer appropriate material to test the system's robustness. We plan to continue validation of NoteAid using other types of notes, including outpatient notes. We evaluated our system by using physician reviews to get feedbacks on the quality of lay definitions. Physicians, who have received high education, might have better computer skills than people who rarely used computers. We will include patients with mixed levels of computer skills in our patient study in the future.

Conclusions

Physician evaluation yielded positive results and useful feedback for content validation and refinement of this innovative tool. We have improved NoteAid based on physicians' feedback. Next steps include a study engaging patients to test the system. Tools such as NoteAid may have the potential to improve patient EHR comprehension, which, when used concurrently with patient portals such as VA's MyHealtheVet [7,54], can improve patient experience, engagement, and health knowledge.

Acknowledgments

This work was supported by the Institutional National Research Service Award (T32) 5T32HL120823-02 from the National Institutes of Health and the Health Services Research and Development of the US Department of Veterans Affairs Investigator Initiated Research 1I01HX001457-01. The content is solely the responsibility of the authors and does not necessarily represent the official views of NIH, the US Department of Veterans Affairs, or the US Government.

The authors would like to thank Sarah Cutrona, Edgard Granillo, Andrew Hsu, Richeek Pradhan, and Renda Wiener for providing valuable feedback on improving the NoteAid system and Weisong Liu and Elaine Freund for technical support in data collection. They also thank the anonymous reviewers for their constructive comments and suggestions.

Authors' Contributions

HY and JC designed the study. TH provided important feedback on study design. JC and ED collected the EHR data. JC and BPR developed the NoteAid system. ED interviewed physicians. ED and JC analyzed the interview data. JC drafted the manuscript. ED contributed substantially to paper writing. HY, CB, TM, DMZ, VGV, and SM provided valuable feedback on system improvement. All authors contributed to the writing and revision of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Criteria for judging importance of candidate terms extracted from electronic health record (EHR) corpora.

[PDF File (Adobe PDF File), 551KB - jmir_v20i1e26_app1.pdf]

Multimedia Appendix 2

Auxiliary data used by NoteAid for medical concept extraction.

[PDF File (Adobe PDF File), 471KB - jmir_v20i1e26_app2.pdf]

Multimedia Appendix 3

Physician responses to open-ended questions in postsession questionnaire.

[PDF File (Adobe PDF File), 416KB - jmir_v20i1e26_app3.pdf]

Multimedia Appendix 4

Coding scheme for qualitative content analysis.

[PDF File (Adobe PDF File), 310KB - jmir_v20i1e26_app4.pdf]

Multimedia Appendix 5

Lay definitions output by NoteAid for medical terms in an electronic health record note excerpt.

[PDF File (Adobe PDF File), 436KB - jmir_v20i1e26_app5.pdf]

Multimedia Appendix 6

User interfaces of the original NoteAid system used in this study and the system with postevaluation enhancements.

[PDF File (Adobe PDF File), 608KB - jmir_v20i1e26_app6.pdf]

References

1. Davis K, Schoenbaum SC, Audet A. A 2020 vision of patient-centered primary care. J Gen Intern Med 2005 Oct;20(10):953-957 [FREE Full text] [doi: 10.1111/j.1525-1497.2005.0178.x] [Medline: 16191145]
2. Detmer D, Bloomrosen M, Raymond B, Tang P. Integrated personal health records: transformative tools for consumer-centric care. BMC Med Inform Decis Mak 2008;8:45 [FREE Full text] [doi: 10.1186/1472-6947-8-45] [Medline: 18837999]
3. Steinbrook R. Health care and the American Recovery and Reinvestment Act. N Engl J Med 2009 Mar 12;360(11):1057-1060. [doi: 10.1056/NEJMp0900665] [Medline: 19224738]
4. Wright A, Feblowitz J, Samal L, McCoy AB, Sittig DF. The Medicare Electronic Health Record Incentive Program: provider performance on core and menu measures. Health Serv Res 2014 Feb;49(1 Pt 2):325-346 [FREE Full text] [doi: 10.1111/1475-6773.12134] [Medline: 24359554]
5. 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]
6. Delbanco T, Walker J, Darer JD, Elmore JG, Feldman HJ, Leveille SG, et al. Open notes: doctors and patients signing on. Ann Intern Med 2010 Jul 20;153(2):121-125. [doi: 10.7326/0003-4819-153-2-201007200-00008] [Medline: 20643992]
7. Turvey C, Klein D, Fix G, Hogan TP, Woods S, Simon SR, et al. Blue Button use by patients to access and share health record information using the Department of Veterans Affairs' online patient portal. J Am Med Inform Assoc 2014;21(4):657-663. [doi: 10.1136/amiajnl-2014-002723] [Medline: 24740865]
8. Delbanco T, Walker J, Bell SK, Darer JD, Elmore JG, Farag N, et al. Inviting patients to read their doctors' notes: a quasi-experimental study and a look ahead. Ann Intern Med 2012 Oct 2;157(7):461-470 [FREE Full text] [doi: 10.7326/0003-4819-157-7-201210020-00002] [Medline: 23027317]
9. Nazi KM, Hogan TP, McInnes DK, Woods SS, Graham G. Evaluating patient access to Electronic Health Records: results from a survey of veterans. Med Care 2013 Mar;51(3 Suppl 1):S52-S56. [doi: 10.1097/MLR.0b013e31827808db] [Medline: 23407012]
10. Woods SS, Schwartz E, Tuepker A, Press NA, Nazi KM, Turvey CL, et al. Patient experiences with full electronic access to health records and clinical notes through the My HealtheVet Personal Health Record Pilot: qualitative study. J Med Internet Res 2013;15(3):e65 [FREE Full text] [doi: 10.2196/jmir.2356] [Medline: 23535584]

11. Nazi KM, Turvey CL, Klein DM, Hogan TP, Woods SS. VA OpenNotes: exploring the experiences of early patient adopters with access to clinical notes. J Am Med Inform Assoc 2015 Mar;22(2):380-389. [doi: 10.1136/amiajnl-2014-003144] [Medline: 25352570]
12. Mossanen M, True LD, Wright JL, Vakar-Lopez F, Lavallee D, Gore JL. Surgical pathology and the patient: a systematic review evaluating the primary audience of pathology reports. Hum Pathol 2014 Nov;45(11):2192-2201. [doi: 10.1016/j.humpath.2014.07.008] [Medline: 25149550]
13. Pyper C, Amery J, Watson M, Crook C. Patients' experiences when accessing their on-line electronic patient records in primary care. Br J Gen Pract 2004 Jan;54(498):38-43 [FREE Full text] [Medline: 14965405]
14. Keselman A, Slaughter L, Smith CA, Kim H, Divita G, Browne A, et al. Towards consumer-friendly PHRs: patients' experience with reviewing their health records. AMIA Annu Symp Proc 2007:399-403 [FREE Full text] [Medline: 18693866]
15. 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]
16. Lerner EB, Jehle DV, Janicke DM, Moscati RM. Medical communication: do our patients understand? Am J Emerg Med 2000 Nov;18(7):764-766. [doi: 10.1053/ajem.2000.18040] [Medline: 11103725]
17. Jones RB, McGhee SM, McGhee D. Patient on-line access to medical records in general practice. Health Bull (Edinb) 1992 Mar;50(2):143-150. [Medline: 1517087]
18. Baldry M, Cheal C, Fisher B, Gillett M, Huet V. Giving patients their own records in general practice: experience of patients and staff. Br Med J (Clin Res Ed) 1986 Mar 1;292(6520):596-598 [FREE Full text] [Medline: 3081187]
19. NCES. 2006. The Health Literacy of America's Adults: Results from the 2003 National Assessment of Adult Literacy URL: http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2006483 [accessed 2016-07-15] [WebCite Cache ID 6j1MrFZFj]
20. 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]
21. 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]
22. Tieu L, Sarkar U, Schillinger D, Ralston JD, Ratanawongsa N, Pasick R, et al. Barriers and facilitators to online portal use among patients and caregivers in a safety net health care system: a qualitative study. J Med Internet Res 2015;17(12):e275 [FREE Full text] [doi: 10.2196/jmir.4847] [Medline: 26681155]
23. Keselman A, Smith CA. A classification of errors in lay comprehension of medical documents. J Biomed Inform 2012 Dec;45(6):1151-1163 [FREE Full text] [doi: 10.1016/j.jbi.2012.07.012] [Medline: 22925723]
24. Wernick M, Hale P, Anticich N, Busch S, Merriman L, King B, et al. A randomised crossover trial of minimising medical terminology in secondary care correspondence in patients with chronic health conditions: impact on understanding and patient reported outcomes. Intern Med J 2016 May;46(5):596-601. [doi: 10.1111/imj.13062] [Medline: 26968750]
25. McCray AT. Promoting health literacy. J Am Med Inform Assoc 2005;12(2):152-163 [FREE Full text] [doi: 10.1197/jamia.M1687] [Medline: 15561782]
26. Berland GK, Elliott MN, Morales LS, Algazy JI, Kravitz RL, Broder MS, et al. Health information on the Internet: accessibility, quality, and readability in English and Spanish. J Am Med Assoc 2001;285(20):2612-2621 [FREE Full text] [Medline: 11368735]
27. Sbaffi L, Rowley J. Trust and credibility in web-based health information: a review and agenda for future research. J Med Internet Res 2017 Jun 19;19(6):e218 [FREE Full text] [doi: 10.2196/jmir.7579] [Medline: 28630033]
28. Graber MA, Roller CM, Kaeble B. Readability levels of patient education material on the World Wide Web. J Fam Pract 1999 Jan;48(1):58-61. [Medline: 9934385]
29. Kasabwala K, Agarwal N, Hansberry DR, Baredes S, Eloy JA. Readability assessment of patient education materials from the American Academy of Otolaryngology--Head and Neck Surgery Foundation. Otolaryngol Head Neck Surg 2012 Sep;147(3):466-471. [doi: 10.1177/0194599812442783] [Medline: 22473833]
30. Zeng-Treitler Q, Goryachev S, Kim H, Keselman A, Rosendale D. Making texts in electronic health records comprehensible to consumers: a prototype translator. AMIA Annu Symp Proc 2007:846-850 [FREE Full text] [Medline: 18693956]
31. Kandula S, Curtis D, Zeng-Treitler Q. A semantic and syntactic text simplification tool for health content. AMIA Annu Symp Proc 2010;2010:366-370 [FREE Full text] [Medline: 21347002]
32. Polepalli RB, Houston T, Brandt C, Fang H, Yu H. Improving patients' electronic health record comprehension with NoteAid. Stud Health Technol Inform 2013;192:714-718. [Medline: 23920650]
33. Chen J, Jagannatha AN, Fodeh SJ, Yu H. Ranking medical terms to support expansion of lay language resources for patient comprehension of electronic health record notes: adapted distant supervision approach. JMIR Med Inform 2017 Oct 31;5(4):e42 [FREE Full text] [doi: 10.2196/medinform.8531] [Medline: 29089288]
34. Zeng QT, Tse T. Exploring and developing consumer health vocabularies. J Am Med Inform Assoc 2006;13(1):24-29 [FREE Full text] [doi: 10.1197/jamia.M1761] [Medline: 16221948]

35. Mowery D, Wiebe J, Visweswaran S, Harkema H, Chapman WW. Building an automated SOAP classifier for emergency department reports. J Biomed Inform 2012 Feb;45(1):71-81 [FREE Full text] [doi: 10.1016/j.jbi.2011.08.020] [Medline: 21925286]
36. Jagannatha A, Chen J, Yu H. Mining and Ranking Biomedical Synonym Candidates from Wikipedia. 2015 Presented at: The Sixth International Workshop on Health Text Mining and Information Analysis (Louhi); September 17-21, 2015; Lisbon, Portugal p. 142-151 URL: https://www.cs.cmu.edu/~ark/EMNLP-2015/proceedings/LOUHI/LOUHI-2015.pdf#page=154
37. Mikolov T, Sutskever I, Chen K, Corrado G, Dean J. Distributed representations of words and phrases and their compositionality. Presented at: Advances in Neural Information Processing Systems 2013 (NIPS 2013); December 5-10, 2013; Lake Tahoe, NV p. 3111-3119 URL: https://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf
38. Attar R, Fraenkel AS. Local feedback in full-text retrieval systems. J ACM 1977;24(3):397-417. [doi: 10.1145/322017.322021]
39. Aronson AR, Lang F. An overview of MetaMap: historical perspective and recent advances. J Am Med Inform Assoc 2010;17(3):229-236 [FREE Full text] [doi: 10.1136/jamia.2009.002733] [Medline: 20442139]
40. Macefield R. How to specify the participant group size for usability studies: a practitioner's guide. J Usability Stud 2009;5(1):34-45.
41. Virzi RA. Refining the test phase of usability evaluation: how many subjects is enough? Hum Factors 2016 Nov 23;34(4):457-468. [doi: 10.1177/001872089203400407]
42. Nielsen J, Landauer T. A mathematical model of the finding of usability problems. 1993 Presented at: The INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems; April 24 - 29, 1993; Amsterdam, The Netherlands p. 206-213. [doi: 10.1145/169059.169166]
43. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput 2003 Aug;35(3):379-383. [Medline: 14587545]
44. Hwang W, Salvendy G. Number of people required for usability evaluation. Commun ACM 2010 May 01;53(5):130. [doi: 10.1145/1735223.1735255]
45. Yu H, Kaufman D. A cognitive evaluation of four online search engines for answering definitional questions posed by physicians. Pac Symp Biocomput 2007:328-339 [FREE Full text] [Medline: 17990503]
46. Zheng K, Vydiswaran VG, Liu Y, Wang Y, Stubbs A, Uzuner Ö, et al. Ease of adoption of clinical natural language processing software: an evaluation of five systems. J Biomed Inform 2015 Dec;58 Suppl:S189-S196 [FREE Full text] [doi: 10.1016/j.jbi.2015.07.008] [Medline: 26210361]
47. Hirsch JS, Tanenbaum JS, Lipsky GS, Liu C, Schmitz E, Hashorva D, et al. HARVEST, a longitudinal patient record summarizer. J Am Med Inform Assoc 2015 Mar;22(2):263-274 [FREE Full text] [doi: 10.1136/amiajnl-2014-002945] [Medline: 25352564]
48. 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]
49. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
50. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: implications for conducting a qualitative descriptive study. Nurs Health Sci 2013 Sep;15(3):398-405. [doi: 10.1111/nhs.12048] [Medline: 23480423]
51. Meystre SM, Savova GK, Kipper-Schuler KC, Hurdle JF. Extracting information from textual documents in the electronic health record: a review of recent research. Yearb Med Inform 2008:128-144. [Medline: 18660887]
52. Xu H, Stetson PD, Friedman C. A study of abbreviations in clinical notes. AMIA Annu Symp Proc 2007 Oct 11:821-825 [FREE Full text] [Medline: 18693951]
53. Liu H, Lussier YA, Friedman C. Disambiguating ambiguous biomedical terms in biomedical narrative text: an unsupervised method. J Biomed Inform 2001 Aug;34(4):249-261 [FREE Full text] [doi: 10.1006/jbin.2001.1023] [Medline: 11977807]
54. Nazi KM, Woods SS, Woods SS. MyHealtheVet PHR: a description of users and patient portal use. AMIA Annu Symp Proc 2008:1182. [Medline: 18999142]

Abbreviations

ADS: adapted distant supervision
CHV: consumer health vocabulary
EHR: electronic health record
ICU: intensive care unit
MFS: most frequent sense
NLP: natural language processing
UMLS: Unified Medical Language System
VA: Veterans Affairs

VHA: Veterans Health Administration

Edited by G Eysenbach; submitted 08.08.17; peer-reviewed by D Klein, B Schreiweis, M Stellefson; comments to author 28.09.17; revised version received 21.11.17; accepted 06.12.17; published 22.01.18.

Please cite as:
Chen J, Druhl E, Polepalli Ramesh B, Houston TK, Brandt CA, Zulman DM, Vimalananda VG, Malkani S, Yu H
A Natural Language Processing System That Links Medical Terms in Electronic Health Record Notes to Lay Definitions: System Development Using Physician Reviews
J Med Internet Res 2018;20(1):e26
URL: http://www.jmir.org/2018/1/e26/
doi:10.2196/jmir.8669
PMID:29358159

Original Paper

eHealth Technologies, Multimorbidity, and the Office Visit: Qualitative Interview Study on the Perspectives of Physicians and Nurses

Graham G Macdonald[1,2], BA, MA; Anne F Townsend[3], PhD; Paul Adam[4], MSW; Linda C Li[1,5], PhD, PT, MSc; Sheila Kerr[6], BPT; Michael McDonald[7], PhD; Catherine L Backman[1,8], PhD, FCAOT

[1]Arthritis Research Canada, Richmond, BC, Canada
[2]Rehabilitation Sciences, University of British Columbia, Vancouver, BC, Canada
[3]Psychology Applied to Health, University of Exeter Medical School, University of Exeter, Exeter, United Kingdom
[4]Vancouver Coastal Health, Vancouver, BC, Canada
[5]Department of Physical Therapy, University of British Columbia, Vancouver, BC, Canada
[6]Arthritis Patient Advisory Board, Arthritis Research Canada, Richmond, BC, Canada
[7]W Maurice Young Centre for Applied Ethics, School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada
[8]Department of Occupational Science and Occupational Therapy, University of British Columbia, Vancouver, BC, Canada

Corresponding Author:
Graham G Macdonald, BA, MA
Arthritis Research Canada
5591 No 3 Road
Richmond, BC,
Canada
Phone: 1 6046794405
Email: gmacdonald@arthritisresearch.ca

Abstract

Background: eHealth is a broad term referring to the application of information and communication technologies in the health sector, ranging from health records to telemedicine and multiple forms of health education and digital tools. By providing increased and anytime access to information, opportunities to exchange experiences with others, and self-management support, eHealth has been heralded as transformational. It has created a group of informed, engaged, and empowered *patients as partners*, equipped to take part in shared decision making and effectively self-manage chronic illness. Less attention has been given to health care professionals' (HCPs) experiences of the role of eHealth in patient encounters.

Objective: The objective of this study was to examine HCPs' perspectives on how eHealth affects their relationships with patients living with multiple chronic conditions, as well as its ethical and practical ramifications.

Methods: We interviewed HCPs about their experiences with eHealth and its impact on the office visit. Eligible participants needed to report a caseload of ≥25% of patients with multimorbidity to address issues of managing complex chronic conditions and coordination of care. We used a semistructured discussion guide for in-depth interviews, and follow-up interviews served to clarify and expand upon initial discussions. Constant comparisons and a narrative approach guided the analyses, and a relational ethics conceptual lens was applied to the data to identify emergent themes.

Results: A total of 12 physicians and nurses (6 male, 6 female; median years of practice=13) participated. eHealth tools most frequently described were Web-based educational resources for patients and Web-based resources for HCPs such as curated scientific summaries on diagnostic criteria, clinical therapies, and dosage calculators. Analysis centered on a grand theme of *the two-way conversation* between HCPs and patients, which addresses a general recentering of the ethical relationship between HCPs and patients around engagement. Subthemes explain the evolution of the two-way conversation, and having, using, and supporting the two-way conversation with patients, primarily as this relates to achieving adherence and health outcomes.

Conclusions: Emerging ethical concerns were related to the ambiguity of the ideal of empowered patients and the ways in which health professionals described enacting those ideals in practice, showing how the cultural shift toward truly mutually respectful and collaborative practice is in transition. HCPs aim to act in the best interests of their patients; the challenge is to benefit from emergent technologies that may enhance patient-HCP interactions and effective care, while abiding by regulations, dealing with the strictures of the technology itself, and managing changing demands on their time.

(J Med Internet Res 2018;20(1):e31) doi:10.2196/jmir.8983

KEYWORDS

eHealth; patient-physician relationship; relational ethics; multimorbidity; online information seeking; email; office visit; health professional perspective

Introduction

eHealth is a broad term referring to the application of information and communication technologies in the health sector, ranging from health records to telemedicine and multiple forms of health education, support, and tools [1]. It has been described as a general orientation toward exploring the possibilities of information and communication technology in health [2]. By providing increased and anytime access to information, opportunities to exchange experiences with others, and self-management support, eHealth has been heralded as transformational. It has the potential to create informed, engaged, and empowered "patients as partners," equipped to take part in shared decision making and effectively self-manage chronic illness [3].

eHealth has been championed as a way to coordinate care among professionals and advance the move away from a disease-centered model of care toward patient-centered approaches better suited to address the needs of patients with multiple chronic conditions [4]. It is thought that eHealth can correct for some of the shortcomings of the health care system that have frustrated patients and caregivers. These include long wait times for appointments and limited access to health care professionals (HCPs), poor communication with and between HCPs, and the challenges that come with managing multiple health conditions (multimorbidity) [5]. However, the changes that various eHealth technologies bring to relationships—between patients and HCPs or between HCPs—have received comparatively little attention in the midst of the digital transformation [6].

Relationships between patients and HCPs are built through communication, and the way and means by which this communication occurs is in the midst of an upheaval precipitated by digital and information technologies. It has been postulated that eHealth technologies have the potential to improve communication between patients (those with multimorbidities in particular) and their health professionals [7]. For example, in a study of patient-physician communication through a Web-based eHealth portal, patients had the opportunity to create and share narratives outside of the allotted time of their consultations, and these narratives helped physicians develop a better understanding of their patients' situations [7]. In a study on the use of a digitally mediated personal physician presence online, it was found that both patients and HCPs welcomed using the platform and benefited from Web-based interaction [8]. Lygidakis et al identified that HCPs' perceived barriers to use were usually dispelled upon becoming familiar with the tools and suggest that training is key to addressing this issue [8].

It has been found that patients' Web-based information seeking can have positive impacts on the relationship with their physicians [9], but only if the accuracy and quality of health information are good [10]. Similarly, Laugesen et al noted that high-quality information does not itself replace the need for a previous trusting relationship with the HCP [11]. Despite these benefits, a 2017 review of 41 papers and 2 chapters published in 2000-2016 regarding adoption of eHealth information and apps notes that "Mainstream medical practice has not yet adapted to the ubiquitous use of the internet by patients" [12]. There remains a need to better understand perspectives from HCPs regarding eHealth because virtually all patient-provider interaction and health information will eventually be mediated by eHealth technologies [13]. It is, therefore, important to examine the new ethical and practical concerns brought about by changes in the patient-provider relationship that technological advancement has already put in motion. In the previous phase of our project, we found that both patients and HCPs were aware of a changing dynamic in their consultations where patients were more informed but uncertain, leading them to have more questions for their HCPs [14]. In general, informed and engaged patients were less reliant on their HCPs and no longer saw them as the gatekeepers of knowledge, having begun to regard themselves as the experts on their condition. Although some HCP participants relayed benefits and anticipated a positive future as more eHealth tools were more widely used, others were reticent to adopt new technology. Findings suggested a need for a clearer understanding of why some HCPs were wary of changing expectations, responsibilities, and obligations arising from patients' use of Internet, apps, and other eHealth tools. The objective of this paper was to examine how eHealth affects patient-HCP relationships, with particular attention on the office visit. We consider ethical and practical considerations from the perspectives of physicians and nurses.

Methods

Design

This analysis is part of a larger, two-phase qualitative study; the protocol is described in detail elsewhere [3]. Ethical review and approval were provided by the University of British Columbia Behavioral Ethics Review Board. The design for the overall study was informed by narrative [15] and constructivist grounded theory [16] approaches, using a relational ethics lens. Relational ethics addresses the ethical content and decisions implicit in everyday relationships and conversations [17]. This suited the overall study goal: to develop an understanding of how patients living with arthritis and multiple chronic conditions and health care professionals perceived the influence of eHealth technologies in managing chronic illness. Findings from the first phase using focus groups with patients and HCPs [14] informed a second phase of in-depth interviews. Here, we focus on the interviews with physicians and nurses about their experiences with eHealth and its impact on the office visit.

Eligibility and Recruitment

Eligible participants were health professionals with at least 2 years of experience beyond their professional degree. They needed to report a caseload of ≥25% of patients with multimorbidity to address issues of managing complex chronic conditions and coordination of care. Participants were recruited by circulating notices to local clinics and medical rounds and by asking colleagues to share notices (word-of-mouth). Generalist and specialist physicians and nurses were purposively targeted, given their small numbers in the phase one (focus group) portion of the main study, so that we could explore the office visit in greater depth. Because the main study focused on patients with arthritis and at least one additional chronic condition, recruitment of specialist practitioners was directed to rheumatology and internal medicine.

Interviews

We used a flexible, semistructured discussion guide for in-depth personal interviews, and follow-up telephone interviews served to clarify and expand upon initial discussions. The discussion guide was informed by findings from focus group discussions with patients and health professionals [3]. Interviews were primarily conducted by a sociologist member of the research team (AFT) experienced in qualitative health research; a research assistant trained in qualitative data gathering conducted interviews and follow-up calls with 3 participants to accommodate HCP schedules. The interview topic guide is appended (see Multimedia Appendix 1), but it should be noted that interviews were conversational in nature and items were not asked verbatim or in the order presented. Audiotapes were transcribed verbatim by a professional transcriptionist and identifying information (names, places) was removed to protect anonymity. Participants are referred to by pseudonyms.

Data Analysis

A narrative approach [18] was taken to ensure that the rich and complex data of the individual interviews were understood and presented in a way that remained faithful to the narrative of the interview and not taken out of context. Using constant comparison method, transcripts of the interviews were read repeatedly by 2 authors (GGM and CLB) to achieve immersion and develop a sense of the whole [19]. This was followed by multiple close readings of the text to pick out keywords and ideas to be used as codes for further organizing the data, a process conducted by the lead author (GGM). This initial analysis documented first impressions and thoughts to begin developing a coding scheme to be applied more broadly to the data, supported by using QSR NVivo 10 software. Once the codes had been applied to transcripts, the coded material was sorted into categories based on how the different codes related to one another [20]. The process of coding and sorting codes into meaningful categories required multiple iterations to adequately explore the different possibilities within the data. As categories were developed by the lead author, they were discussed with the last author until agreement was reached on labeling key categories. To minimize bias and achieve clarity, preliminary codes and transcript excerpts were shared with team members to test impressions and assumptions and revise the coding framework. The categories were then analyzed through a relational ethics lens to identify emerging themes along the lines of what was being valued and what was at risk in the relationships described by participants [21]. Of the 7 research team members, 4 read all transcripts and 3 reviewed excerpts selected by the primary analysts (GGM and CLB) based on their relevance to the emerging themes. The team discussed iterations during the analytical process by email, in person, or by phone in scheduled team meetings, and by responding to draft reports of findings. Collectively, the team brought perspectives from patients, practitioners, and sociology, health services, and ethics researchers.

The final analysis is represented by a grand narrative (the two-way conversation, comprising 4 categories) and 2 small supplementary narratives related to email conversations and anticipating the future of eHealth. Finally, quotes were selected to illustrate and substantiate the final narratives shared here. Pseudonyms are used for anonymity.

Results

A total of 12 HCPs participated, each giving an in-depth interview of approximately an hour (range 25-78 min, mean 51 min; 9 in person, 3 by telephone) and a 12 to 31-min (mean 23 min) follow-up interview by telephone. All participants practiced in urban settings, with some serving patients from rural and remote populations who traveled to the city to see them. The HCPs worked in hospital, clinic, and private office settings. Of the participants, 3 were general practitioners, 2 were registered nurses, 1 was a nurse practitioner, 3 were rheumatologists, 1 was a physician clinician-scientist, and 2 were rheumatology fellows. Out of the total 12 HCPs, 6 were male and 6 were female, the median age was 46 years, and their years of practice ranged from the final year of rheumatology fellowship to 29 years, with a median of 13 years (Table 1).

Participants mostly used eHealth information resources during their office visits, with a few other technologies less frequently (Table 2).

Web-based resources for patients were most common, with all participants using this in some form or another. A majority of participants (n=10) used Web-based information sharing or education sites for physicians (such as UpToDate.com), some used email to communicate with patients, and some used electronic record systems to which patients may or may not have access. A third of participants used digital diagnostic tools (apps or online) in their office visits with patients. An example of how eHealth tools and resources were accessed and used is Donna’s description of a professionally curated online site that she finds useful in her practice as a rheumatologist, as shown in Textbox 1.

Table 1. Participant characteristics.

Pseudonym	Profession	Gender	Years in practice
Arlene	Nurse	Female	27
Donna	Fellow	Female	2
Elise	Nurse	Female	25
Gabriel	Physician clinician-scientist	Male	29
Henry	Fellow	Male	<1
James	Family physician	Male	13
Jocelyn	Rheumatologist	Female	25
Luanne	Rheumatologist	Female	6
Martin	Family physician	Male	2
Miranda	Rheumatology nurse	Female	18
Peter	Family physician	Male	<5
Richard	Fellow	Male	Not provided

Table 2. eHealth tools and technology cited by participants.

Tool or technology	Used and discussed	Mentioned, not used
Monitoring apps, for example, to monitor blood pressure, weight, and physical activity	1	4
Web-based information/resources for patients	12	-
Web-based resources (curated sites) for health care professionals (HCPs), for example, diagnostic criteria, disease activity calculators, and medication dosage calculators	10	-
Web-based portals with patients to access personal medical records and test results	5	-
Videoconferencing for patient visits and telehealth consultations	1	4
Email with patients	5	-
Social media	4	7
Diagnostic tools/decision aids	4	1

Textbox 1. Example of how eHealth tools and resources were accessed (Donna’s description of a professionally curated online site).

Donna: So a company has to pay to access and then essentially they’ve made a whole bunch of articles spanning all areas of medicine that are kept, as it says, “up to date.” So every 3 to 6 months someone goes in, looks at all the literature, and updates it. And they get people that are experts in that area to keep it updated. And so it’s an easy way to go to without having to go through all the literature yourself or go to your textbooks, which are now out of date to get a quick answer.

So if you are in the clinic and you’ve just forgotten the appropriate treatment for someone, you can look it up, which is handy when you’re still in the learning process. And it breaks things down into talks about the causes for the disease, different things to think about, all the medications, how you treat it, pros and cons, limitations of our knowledge, how to follow people.

It also has a drug calculator and it can tell you, you know, all the background information about medications, what the dose is if they have, say, an organ problem like kidney failure, if you have to dose adjust for that or whether or not you have to. It can tell you that. It tells you what to monitor, what to watch out for, and when not to use the medication. And you can also do a drug interaction. So if someone is on, for example, one of our medications, allopurinol, it often interacts with other medications. So you can put that in and then check to make sure there aren’t interactions before you prescribe it.

The Two-Way Conversation

The underlying theme identified was labeled “the two-way conversation.” It concerns HCP-patient communication, describing the dynamics of a conversation that is in the process of change (regardless of how long the participant had practiced). Several participants described this as a shift toward a two-way conversation in medical visits, explaining how a more collaborative interaction with patients has evolved, in part facilitated by eHealth technologies.

We first explore the HCP perspective on the *evolution* of this two-way conversation through the rise of eHealth, and then examine their perspectives of the impact of eHealth on the present state of collaborative consultation: *having*, *using*, and *supporting* the conversation with patients. We present ways in which the two-way conversation is seen to facilitate the therapeutic relationship: how it is viewed as helpful to HCPs and how they try to support patients to be better able to engage in a two-way conversation.

Evolution of the Two-Way Conversation

Participants described the historical relationship between patients and HCPs as having changed socially, technologically, and generationally from a relationship defined by a largely paternalistic power dynamic toward one based on collaborative partnership (Donna—Textbox 2).

They attributed this to the rise of the Internet bringing about widespread availability of information, meaning that patients were no longer relying solely on their HCPs for all of their information regarding their health (Luanne and Elise—Textbox 2). The Internet has facilitated this shift through providing the platform for patients to organize themselves in online communities where they can discuss their conditions, share favored resources, complain, or offer support to one another, connecting and empowering them (Jocelyn—Textbox 2). A characteristic of this more cooperative model is that more informed patients are more engaged and involved in making health decisions with their HCPs (Elise and Jocelyn—Textbox 2). These broad changes were regarded by participants as a generational shift in attitudes and ideas around health care practice among both patients and HCPs, to which HCPs of previous generations may have difficulty adapting (Peter—Textbox 2).

Having a Two-Way Conversation

Overall, participants were enthusiastic about more informed patients as contributors to better health outcomes. There were a range of approaches to two-way dialogues described by participants, from "partnership," emphasizing collaboration and teaching, to "alliance," emphasizing patient choice. Martin embraced the idea of patients as "partners," seeing a partner as someone "who simply helps me in bettering an outcome" by educating themselves and conscientiously monitoring their condition and behavior (Textbox 3).

Peter echoed this notion, stating frankly that patients who are engaged through eHealth and informed about their condition "are more useful, clinically" (Textbox 3). Peter related the usefulness of informed patients to their understanding clinical language that enhanced their ability to hold a "back-and-forth conversation" as opposed to "a one-way conversation where you would tell them" (Textbox 3).

The two-way conversation is more difficult when there are opposing views being expressed. Henry explains how he navigates these discussions, emphasizing the importance of maintaining the "therapeutic alliance" and treating patients as adults who "can make up their own minds" (Textbox 3). Gabriel (Textbox 3) outlined a similar approach, offering to discuss the scientific merits of different treatments with patients who came to him interested in trying alternative or non-Western medicine that they discovered on the Internet. Throughout their interviews, Henry and Gabriel used more paternalistic language to describe their interactions with patients than most participants, that is, suggesting or stating they knew what was in the patient's best interest, but still stressed the importance of "letting" patients make their own decisions. They identified their medical expertise as the reason patients seek care, and subsequently used very direct terms to share their opinions. Henry and Gabriel took a hands-off, laissez-faire approach to patient education that placed greater importance on patient autonomy and less emphasis on the pedagogical role of the HCP.

Textbox 2. Evolution of the two-way conversation.

Donna: Historically...physicians have been seen as someone not to question. [Patients from an older generation have] kind of come in and been more passive, in a sense, because theirs is more the paternalistic model, right, which is to do what the doctor says. Whereas the newer generation, it's more as it should be: a collaborative effort.

Luanne: ...the whole paternalistic pattern is changing toward more patient centered and then more patient driven. But it has been facilitated by the Internet and the availability of information.

Elise: But patients were almost 100% reliant on what their physician or what their health care provider tells them. That was their primary source of information. And so now it's very much more a team working together because they are coming with the information that they have sought from all sorts of resources and it's usually the Internet. And, you know, sometimes they might be more up to date with the most recent research than you might be, for example, as a health care provider. So it really does become a team. And patients now, they don't need to rely on their physician to give them the information they can find themselves. And I think it's huge in that they can then make, and they are making, a lot more decisions for themselves about their health care.

Jocelyn: ...the Internet is what I think has enabled that change for a number of reasons. I think it allows people to not feel alone, to be able to more easily connect with other patients...It's sort of like the strength in feeling like you're part of a community as opposed to be isolated that I think empowers people to be able to do that. I think the Internet provides them with knowledge so that they're not relying on their physician for the knowledge...There's more than just knowledge. There are tools on the Internet which some people use and some people don't that are empowering. So I think it empowers patients to make them arrive at the table more informed than in the relationship of dependence.

Peter: I think I'm part of a generation that we have kind of learned, we've been brought up with all this technology and so I think we, I'm not trying to put myself on a pedestal, but like I feel like we are much more equipped at using these tools as opposed to even my mentors who started practicing 30 years ago. I think it's a lot easier for us to use this. Even if somebody, you know, a patient comes in throwing tools at you. I may better be able to handle that through the computer. We have more knowledge to discuss that with them as opposed to somebody who was, you know, trained 30 years ago. They might just deny it and get angry by that.

Textbox 3. Having a two-way conversation.

Martin: Well, I want them to be my partner. I think outcomes are better and it's certainly easier for me if the patient can do some of their own education and some of their own monitoring. Well just to help me out, I mean I have more patients than I can handle and so any bit of help is welcome.

Peter: ...it gives [patients] access to health information that they might not have had access to in the past...[they are] able to communicate their concerns and hopes and expectations and connect it with, you know, specific health. They are more useful clinically,...able to communicate what they want in a better, more effective manner with clinicians as opposed to it just being a one-way conversation where you would tell them. They're able to do more of a back and forth conversation now. I guess you'd say that it's empowering them more than it was before.

Henry: Well I think different people have different approaches. You know, at the end of the day, saying something and someone not hearing it and then saying it louder is certainly not like a solution. So if they're sort of committed to a certain framework I think it's important that I discuss the alternative. But, you know, you don't want to damage the therapeutic alliance. So you propose your own perspective and provide your own resources. And beyond that, as long as they understood everything you said, they're adults and they can make up their own minds.

Gabriel: And they went to this XYZ website. Usually it's "What do you think about these natural remedies for this?" And I think that's the ones that they come looking for alternatives that is not traditional Western medicine because they read all the side effects or they're looking for validation of the natural alternative therapies...Well I say that pretty much in medicine that what we do these days is based on science and I'm not against that [alternative]. If they have good science we can discuss that and if it's just based on no science then I can't say. I just say, well you can try it if you want. [But]...the time that you're going to lose before you're getting proper treatment is valuable time.

James: We have to be people who not just have that information but we have the way to interpret it and we're the ones to say, "Well you read about this study online or you read this thing on [X's] website. Well, you know what, this actually hasn't been shown or this was a rat study or there are actually studies that [show] this could be dangerous or this was a study done on post-menopausal women and you're an 18 year old guy. And so this is how it may not be applicable." So it's really exciting because we can actually do more. The person is sort of coming in already thinking about things, forming questions and so I think it really helps us to perform health literacy in a much more meaningful and deeper way. And I think that there are patients that really value that.

James described a pedagogical approach to interacting with the eHealth users among patients (Textbox 3). In this view, the HCP understands "the way to interpret" Web-based information and has the responsibility to impart accuracy and applicability, and encourage patients to improve "health literacy in a deeper, more meaningful way." He conveyed genuine interest and excitement with two-way conversations with informed patients. Other participants expressed similar notions, being enthusiastic about the ability of patients to do the research and assimilate complex information, but also holding that HCPs were the fail-safe for patients who often lacked the skills and tools to critically assess the information. Arlene talked about bringing a breadth and depth of knowledge to the conversation:

> *So what differentiates the really educated patient on a certain thing from someone who has trained in that? I suppose the main thing is that the health care professional is trained in multiple areas. So if someone becomes an expert in their own disease but they don't know that disease and how it is [related] to other diseases.*

Using the Two-Way Conversation for Patient "Buy-In"

Participants spoke of valuing the two-way conversation as an opportunity to obtain "buy-in" from patients. When acknowledging shared decision making, the more prevalent description was of a conversation that focused on providing information or rationale for the recommended treatment. For example, Jocelyn mentioned the necessity of building patient understanding of their health conditions to improve treatment adherence (Textbox 4).

Richard and Donna saw this approach as essential for countering misconceptions about a medication that was frightening due to what the patient perceived as its "poor reputation" (eg, potentially harmful effects or undesirable characteristics) but from the HCP's perspective had solid evidence as effective (Textbox 4). Jocelyn expanded on the topic of adherence to emphasize that empowerment of the patient through increased access to information and use of eHealth tools helps the HCP to build better relationships and understand patients—especially ones with multi-morbidities—more thoroughly and holistically (Textbox 4).

Supporting the Patient to Engage in a Two-Way Conversation

The quality and quantity of information available to patients were the main concerns of all participants. They shared that the ability to have a productive two-way conversation relied on how patients used the Internet, apps, or decision tools.

Because it was easy to be overwhelmed by the amount of information some patients brought to the visit, most participants had developed management strategies. For example, Martin, a family physician, recommended specific websites he considered trusted sources, as did Luanne, a rheumatologist who emphasized the importance of valid websites to patients (Textbox 5).

The 3 nursing participants (Arlene, Elise, and Miranda) saw it as part of their job to spend more time directing patients to resources, to save the physician time. Miranda noted that when patients came in to the office visit informed but not necessarily with accurate information, she viewed her role to "steer them in the right direction" after inquiring about where they found their information (Textbox 5). Arlene initially "steers" patients away from blogs and chat rooms because of their emphasis on negative experiences that she thought might skew patient perceptions (Textbox 5).

Textbox 4. Using the two-way conversation for patient “buy-in.”

Jocelyn: I think that if patients are more engaged in the decision and they understand why they are taking a medication and they bought into the decision, that they will have better adherence. So I think it’s kind of a by-product of that as opposed to a direct thing, it’s not so much that I think that the tool itself targets adherence. But I think that if they are more engaged and they feel more part of it, the main thing about adherence—and that’s what I discuss with them—is if they understand why they are taking a medication, they are way more likely to follow up with the instructions. It’s when they don’t really understand why they are taking it, then as soon as it’s inconvenient or they forget or they have a bit of a side effect then they’re less likely to be adherent.

Richard: [referring to decision tools] And I really liked the idea of education especially for methotrexate and different medications because adherence to therapy is a big issue, I think, with patients. And getting them to buy in on what we’re doing and why we’re doing it—that will definitely improve patient outcomes.

Donna: So yeah, definitely, the next time I come across a patient that really dislikes Methotrexate, I will be referring them [to electronic resource]. And it’s also probably a good thing to have anyways for them to know about and to be able to go use as a reference. Because some patients whom we prescribe medication, they don’t want to tell you. They’ve just met you and they don’t feel comfortable telling you that they don’t quite trust you yet or they don’t trust the plan. So they’ve left and you think that they’ve started on something and then they come back six months later and they haven’t started it. So I think having them go to a site can help educate but also increase the actual adherence to treatment.

Jocelyn: I think it’s more than just adherence to medications. I think if we have an engaged patient then we empower them to be able to really manage their disease a whole lot more efficiently. So I think we’ll have better outcomes because we’ll have better care that overall fits better with the patient. So that to me is, we’ll have a better relationship with the patient, we’ll have better management of things other than just the medications and the medical aspects of the disease, which is really important for the quality of life and the dealing with the person as a whole and the disability of the whole.

Textbox 5. Supporting the patient to engage in a two-way conversation.

Martin: Patients are doing a lot of their own research now. Often they’ll come in and say, “What do you think of this that I read about? What do you think of that?” I must admit that some of the newer things they’re more aware of than I am. I tell them, you know, if you read something that just came out, you probably know more about it than I do. And they’re fine with that. They realize the speed of information. So I just really tell them that there are some reputable sites and there are some not reputable sites. So, you know, if you were on the Berkley site or the Mayo site or VGH site or Health and Welfare Canada, we trust those sites.

Luanne: So depending a little bit on what they’re in for, some of it is patient support. So Arthritis Society I will send them to, I have lupus patients for BC Lupus Society has got a support page. If they just want more general medical information I can also send them to places like the Mayo Clinic, Rheumatology Network as other reliable sources of information. And I try to stress with the patients that whenever they’re going to look for sources of information, they should make sure that it’s coming from a valid provider.

Miranda: I can’t say it’s often that they come in with accurate information. There is so much information and I think how people search is what pops up. And if you check with people briefly about where they found their information, I think that most people come in somewhat informed and then we can kind of steer it down the right path from that point onward or at least help them, guide them down the right path.

Arlene: I tend to try to steer people away from blogs at least initially or those chat rooms because often times the people that are on those are not people that have had positive experiences or are having a positive result. I think it skews their perspective of what is going on. But the power of the personal story is really far more riveting and convincing.

Gabriel: Yeah for me, for instance, the use of sites, I know patients when they come to you and you have to provide information they usually get shocked first to get a diagnosis and second to start treatment. And so I give them readings. I print some information for them and tell them if they have more questions to go to these sites and then you come back with me and we can discuss it if you want.

“Guiding” or “steering” patients toward certain resources was a strategy common to all participants, for purposes of ensuring patients had reliable information and to acknowledge time issues. For example, time can be saved in the visit if patients avoid unreliable websites; as Gabriel notes, recommending reliable Web-based resources for questions that arise outside the office visit helps patient learn about, and possibly emotionally adjust to, a diagnosis or new treatment (Textbox 5).

HCPs were concerned not simply about the quality of the information on the Internet but about the patient experiences while searching for answers. Richard remarked that patients are “often misled by the Internet” and that this “can cause a lot of anxiety.” Elise explained how the quantity of information on the Internet itself could be “incredibly overwhelming” to patients trying to make sense of their situation. The following story from Luanne also points to the fact that HCPs should check their own propensity for overwhelming patients with information and should instead opt for digestible and accessible formats:

> *It’s really simple and when I first saw it, I thought, oh dear patients are going to be offended because it’s a little, dumbed down a little bit. I have yet to have anyone complain about that. They’re all, like, oh a single piece of paper. I can read this. There are pictures. There are not too many words and it’s fantastic. And so I think, you know, knowing that there are Internet sites that will give you that extra information if you need it, I don’t feel bad giving them a single page suggesting places they can go and read more if they want to.*

Although the two-way conversation comprised the central narrative, the analysis generated two smaller narratives related to HCP-patient interactions, as follows.

Conversation Beyond the Consultation: A Role for Email?

Different views were presented regarding potential benefits and drawbacks to using email with patients (Textbox 6).

One of the main concerns was related to time constraints and how email could create imbalance in their workload. Martin remarked that “when it’s an office visit, I feel like there’s a finite amount of time to have a consultation.” There are clear time boundaries in the office visit, he notes, but with email, “it goes on and on.”

Richard believed there should be a “professional barrier” between HCPs and patients in regard to email because an expectation exists that messages can be answered immediately, but the reality remains that this is not possible either due to lack of time or the need for a face-to-face discussion. Richard also suggested that “if a doctor would answer every email from every patient, that’s more than a full-time job.” Jocelyn was wary of setting up expectations with email because of the round-the-clock responsibility that it could entail:

> *They’re expecting me to check and respond because maybe in the past I have responded quickly. I’ve set up expectations. So instead of going to the Emerg or going to a walk-in clinic, they wait for my answer. And then maybe I don’t and it was an urgent situation and you don’t seek care. So that scares me. So it’s the expected immediacy of response I think...I think once you’ve started responding quickly, it sets up an expectation and I wonder to what extent I’m then responsible.*

Likewise, Elise (Textbox 6) expressed how the move to electronic records in her office necessitated that “we’ve had to become even more aware of the privacy regulations” and when it came to email with patients “we actually shouldn’t be doing it” because of privacy issues. James, however, seemed to have worked out a system of regulating his email exchange with patients to enhance rather than impede efficiency. James offered that email (sometimes combined with telehealth approaches) could be the best use of his and his patients’ time, given that some experienced mobility issues, lived in remote locations, or both, and not all consultations required a face-to-face visit to be effective.

Textbox 6. The conversation beyond consultation: a role for email?

Martin: Yeah there’s no end to it really. That’s the problem is, you know, the patient will ask a question. I’ll reply and then they’ll reply. I just don’t, you know, when it’s an office visit, I feel like there’s a finite amount of time to have a consultation. And then, you know, both the patient and I say, “okay, time’s up for this appointment. If you want to talk some more, come again.” But that doesn’t happen with email. With email it’s just, it goes on and on.

Richard: I try not to do that [email] because I think there has to be some professional barrier between patients and their physicians; otherwise sometimes you need a little bit of time for things to evolve. And I can understand if patients are anxious and they want to know the answer now. But sometimes you need time. You can’t answer it now. Blood work takes a week to come back or, you know, this test takes a little bit of time. Sometimes symptoms take awhile before you can really tell is this, what direction are we taking and, you know, if a doctor would answer every email from every patient, that’s more than a full-time job.

Elise: And, you know, since we—in one of the offices I work in—since we’ve gone from paper to electronic charts, we’ve had to become even more aware of the privacy regulations and, you know, I’m not sure how but it was brought to the office’s attention that emailing is, that we actually shouldn’t be doing it…you know, patients love it as it’s so it’s easy for them, right? They don’t have to worry about waiting for someone to answer the phone. You know, they can send it at midnight if they are thinking about something. But it’s really, we shouldn’t be doing it.

James: I mean many, the majority of these consultations are for people with whom I’ve had face-to-face consultations and we’ve made a plan. But to carry out that plan I don’t necessarily need them in the same room. And maybe part of the plan is that they have difficulty getting in to see me because of some of the barriers that we talked about. And for someone to come in and just show me the blood pressure readings that they’ve gotten, that’s not a great use of their time. And if they’re able to email it to me and then we’re able to talk about it, it’s efficient use of their time. It’s efficient use of my time and the patients really enjoy it.

Textbox 7. Looking toward the future of eHealth.

Richard: I think that’s an area of opportunity...I think it would be great if patients had copies of their health records on some app, had copies of their imaging. I think that would be very useful. If we had a centralized system that would be useful. Like having [Hospital A, Hospital B, and Hospital C] in different systems is ludicrous, I think.

James: So I mean like anything health IT is a tool. And it has to be used properly. And if we go back to the analogy of the stethoscope, just because someone has a stethoscope doesn’t mean that they’re going to hear the murmur or they may mistake something to be pneumonia when it’s congestive heart failure. So the mere fact that having a stethoscope doesn’t mean that someone is going to use it properly and make good decisions about it unfortunately. And so the same thing holds here in that someone may use the tool, for example, video teleconferencing and see a patient online in a way that’s not productive or maybe for a complaint that isn’t appropriate or something that does need to be seen in person. So I think it’s important that as we gather experience and see what the limits are, those things will become clear and people will get together and there will be more guidelines. But right now we’re really on the cutting edge of people saying, hmm, okay we have this technology. There’s a potential for abuse but a potential for great benefit. But what we need to do is say, okay, are there guidelines that apply to everyone using it? Are there guidelines that apply to some people using it and some sorts of patients? And what parameters can we put on so that we’re using this to the best of our abilities?

Looking Toward the Future of eHealth

Some participants spoke of the potential of eHealth tools and technology not yet routinely used in their practice. For example, they envisioned improved or new tools as paving the way toward more efficient team care by making coordination between HCPs easier, as in this example from Elise:

> *So many patients come in and they say, "Well I'm on some kind of blood pressure medication and I don't know what it's called." So we can access that now from the office and you can find all of their medications [with patient consent, on the provincial network that records all prescriptions dispensed]. So in that way it's a huge safety, you know, protection for the patient because you then know interactions and things like that…Then if a patient is seeing multiple specialists, which they often are..., they should all have access to everything, right? So it's coming. It's not there yet. We still spend quite a bit of time phoning over to a specialist's office and saying, "Can we get that consult?" Because unless they specifically share it when they're doing their dictation, we're not going to see it, right?*

The HCPs in our study described feeling burdened with the inefficiencies of the still-developing eHealth systems they are working within, in particular the electronic health record. It was declared that adequate resources are not in place, systems are not interoperable, and patient concerns and needs may outstrip what HCPs can provide (Richard— Textbox 7).

As a result, they expressed a need for improvements and innovation to ensure the systems they work within were truly functional. James (Textbox 7) made the case that although there may be many tools available, there is no guarantee that those tools will be used effectively and that guidelines drawn from experience should increase awareness of the capacity of eHealth to support practice.

Discussion

Principal Findings

The HCPs in our study saw eHealth technology as an important contributor to how their relationships with patients have changed and are evolving toward more collaborative care—both collaboration with the patient and with other providers. Participants expressed a desire for guidelines for how patients and HCPs communicate and use eHealth technology to optimize benefits and prevent undue burden or abuse. There were participants who were enthusiastic about emailing patients, and others had serious reservations about this. However, there was agreement that if this mode of communication was to work, it would need to be on the terms the HCP could manage in the context of their practice, which also included adhering to privacy regulations (the mechanisms of which were not always well understood by HCPs). Participants were concerned about how Web-based information, health apps, and email with patients could create unrealistic expectations for patients and unsustainable responsibilities for themselves in terms of workload and time.

Participants said they valued informed and engaged patients and saw patient engagement in health care as facilitated by eHealth, leading to better health outcomes. In general, the emergence of eHealth resources was regarded as a positive progression, despite hiccups and shortcomings. Though HCPs no longer see themselves as the gatekeepers of medical information, they described themselves as best able to translate that information for the benefit of their patients. Indeed, the most universal use and discussion of eHealth among HCPs in this study focused on how patients access health-related information on the Internet, with attention given largely to guiding patients toward higher-quality information and better understanding of that information. In describing their consultations with patients, participants seemed to adopt the discourse of collaboration, engagement, and empowerment. They were clearly caring and compassionate, yet elements of a paternalistic relationship with patients were evident in some descriptions. As outlined below, these findings raise important ethical issues.

The ways in which HCPs described how eHealth facilitated engagement and empowerment of their patients offer a window into a more nuanced understanding of shifting relational dynamics and practice ideology in a period of transition. Importantly, our participants did not view patient engagement or empowerment resulting from the rise of eHealth as eroding or taking away their power as HCPs. The engaged or empowered patient was described as an actor in the relationship who could more meaningfully participate in their care. The "two-way conversation" integrated eHealth tools and Internet information used by patients and HCPs, and was variously described as a way to teach patients about their illnesses, facilitate understanding of and adherence to medical advice, support how patients self-manage their chronic diseases, and enhance clinical effectiveness or efficiency.

Some descriptions of teaching patients evidence residual paternalism in that the two-way conversation, as described by some HCPs, did not denote a symmetrical flow of information, but rather cast the engaged, empowered, knowledgeable, and tech-savvy patient as more receptive to medical advice and open to the medical worldview. Most telling, perhaps, is the insight that HCPs were interested in how supporting patients to be engaged and empowered could facilitate "buy-in" or concordance with the HCPs' perspective and judgment.

From the perspective of HCPs, patients were historically treated as subjects upon whom health care was practiced; the eHealth revolution has accelerated the patient-centered care philosophy and facilitated a relationship wherein the patient is a person who meaningfully participates in the process of his or her own treatment. The emergence of this new relationship between patients and HCPs is, however, far from complete, and it is unclear whether the role of eHealth will be to facilitate further changes in power dynamics or maintain the hierarchy between patients and HCPs [22]. Supporting patient engagement is not an either-or commitment but has different forms along a continuum: from consulting with patients and sharing information with them about a diagnosis, to involvement and asking them about their preferences for their plan, to partnership and sharing leadership and decision making [23]. Our findings suggest that there remains a tendency among some HCPs to view patient engagement and empowerment narrowly as a clinical tool rather than a collaborative endeavor.

Although participants candidly responded to questions, in interpreting these findings, it is reasonable to reflect on the research interviews as a social encounter where participants may feel pressed to respond in certain ways to present as current and knowledgeable professionals. Ambiguity can be detected between some of the ideals expressed and the way practices were described. For example, all participants genuinely conveyed a desire to achieve the best possible health outcomes for their patients and welcomed the notions of patient-centered care, collaboration, and shared decisions, primarily in the form of generic descriptions. Yet, descriptions of actual patient encounters that cite adherence and “steering” patients in the right direction show how challenging it can be to fully develop mutually respectful relationships that connect with the experience of what it is like to be a patient. HCPs’ work was also restrained by the practice environment, creating tension between ways they prefer to practice and ways they had to practice due to suboptimal systems. These observations have practical implications for reflective practice, such as encouraging HCPs to consider their goals for patient-centered care and how their practices, attitudes, and word choice may be (mis)interpreted by patients.

Comparison With Prior Work

A study by Laugesen et al indicating that the quality of Internet health information has some impact on patients’ concordance with their physicians is further borne out by the practical concerns of our participants in guiding their patients to certain Internet resources over others [11]. Our participants also drew the connection between the quality of health information accessed by patients and patient compliance, which agrees with the finding of Laugesen et al that there is an indirect relationship between the two [11]. The changes in patient-HCP communication brought about by eHealth were heralded by participants as opportunities to teach and communicate with patients that could increase quality of care and lead to better outcomes, and taken together with the findings of Haskard-Zolnierek and Di Matteo imply that if eHealth technologies do indeed facilitate better HCP communication, they will likely facilitate better adherence as well [24].

From the standpoint of relational ethics, which locates the practice of ethics in the dynamics of everyday relationships, respect for the lived experience of patients is a basic requirement of ethical engagement [17]. Many of our participants either supported or advocated patient empowerment, but referred to this concept in concert with constructs that implied a hierarchy in decisions about care, such as “adherence.” Adherence was valued because it was associated with better outcomes, from a medical standpoint. Following Ajoulat et al, ethical issues may arise where HCPs too narrowly interpret what empowerment is and means to their patients, leading to other aspects of the illness experience being overlooked [25]. An attendant ethical issue often lies with a blurring of the distinctions between having responsibility and having power. Leveraging patients’ high level of engagement with eHealth to transfer onto them the responsibility for aspects of their health care will by no means necessarily result in their empowerment, as it is entirely possible to have responsibility without having power.

Our findings are in agreement with Kreps and Neuhauser [26], who have noted that eHealth technology, through providing timely, accurate, and accessible information to all stakeholders, can support decision making that meaningfully involves patients and providers, but only if they are interoperable and tailored to engage the personal context of patients. Furthermore, our participants agree that the holistic needs of patients with chronic illness demand technologies that facilitate a person-based rather than disease-based model of care [4]. Tools such as online patient diaries that bring the patient experience closer to the HCP [7,8] have the potential to help create the conditions for office visits to incorporate the relational ethics of person-centered care into practice by creating the space for a more effective meaningful exchange of information and negotiation and support of behavior change between patients and HCPs. Some of the frustrations with systems, such as electronic medical records that were not integrated across institutional boundaries, or worries about being inundated with requests if electronic communications were opened up to patients suggest an element of moral distress. Austin [27] described how environmental constraints contribute to moral distress, such as lack of time or structures to communicate with other team members or engage in problem solving or building relationships. Future research could explicitly examine frustrations as an indicator of moral distress to develop potential solutions to support practice.

Collectively, our participants used a relatively narrow range of eHealth tools and solutions, focusing primarily on health information, literacy, and email. Participants with a broader range of experience may raise different issues. However, based on prior studies [12], they may be typical practitioners and findings may inform health professional education specific to eHealth [12], better equipping HCPs to adopt eHealth technologies to strengthen patient-HCP relationships. Twelve respondents is a relatively small number, but issues were repetitive across participants, even though they had differing years of practice, experiences, and familiarity with eHealth, and differing descriptions of communication and caregiving styles. Thus, responses may be transferable to other physicians and nurses in office and clinic practice environments. Nevertheless, as eHealth technologies continue to evolve, examining their impact on effective patient-HCP relationships and health care decisions warrants investigation using different research designs, with larger, more heterogeneous samples that include populations with low health literacy and limited access to eHealth technologies.

Conclusions

HCPs are at the forefront of dealing with the everyday ethical issues emerging from the growing role that eHealth technologies are playing in health care consultations. Their task is simultaneously to use emergent technologies to enhance their interactions with patients and facilitate a beneficial involvement in their health care, all the while abiding by the regulations of existing health care institutions, dealing with the strictures of the technology itself, and trying to manage the changing demands on their time. Potentially, eHealth supports the evolving nature of the reciprocity of the patient-HCP relationship, toward patient-centered care, enhanced communication, and efficient health service delivery.

Acknowledgments

The authors thank participants for sharing their time and insights; Jenny Leese for assistance with participant recruitment and interviews; Dr Kam Shojania for facilitating participant recruitment; and Gordon Whitehead for early contributions as a patient collaborator to the study design. Funding was provided by an ethics catalyst grant from the Canadian Institutes for Health Research.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview guide.

[PDF File (Adobe PDF File), 86KB - jmir_v20i1e31_app1.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 Feb 24;7(1):e1 [FREE Full text] [doi: 10.2196/jmir.7.1.e1] [Medline: 15829471]
2. Eysenbach G. What is e-health? J Med Internet Res 2001;3(2):E20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
3. Townsend A, Adam P, Li LC, McDonald M, Backman CL. Exploring eHealth ethics and multi-morbidity: protocol for an interview and focus group study of patient and health care provider views and experiences of using digital media for health purposes. JMIR Res Protoc 2013 Oct 17;2(2):e38 [FREE Full text] [doi: 10.2196/resprot.2732] [Medline: 24135260]
4. Zulman DM, Jenchura EC, Cohen DM, Lewis ET, Houston TK, Asch SM. How can eHealth technology address challenges related to multimorbidity? Perspectives from patients with multiple chronic conditions. J Gen Intern Med 2015 Aug;30(8):1063-1070 [FREE Full text] [doi: 10.1007/s11606-015-3222-9] [Medline: 25691239]
5. Gill A, Kuluski K, Jaakkimainen L, Naganathan G, Upshur R, Wodchis WP. “Where do we go from here?” Health system frustrations expressed by patients with multimorbidity, their caregivers and family physicians. Healthc Policy 2014 May;9(4):73-89 [FREE Full text] [Medline: 24973485]
6. Watcher R. The Digital Doctor: Hope, Hype and Harm at the Dawn of Medicine's Computer Age. New York City: McGraw-Hill Education; 2015.
7. Das A, Faxvaag A, Svanæs D. The impact of an eHealth portal on health care professionals' interaction with patients: qualitative study. J Med Internet Res 2015 Nov 24;17(11):e267 [FREE Full text] [doi: 10.2196/jmir.4950] [Medline: 26601678]
8. Lygidakis C, Wallace P, Tersar C, Marcatto F, Ferrante D, Della Vedova R, et al. Download your doctor: implementation of a digitally mediated personal physician presence to enhance patient engagement with a health-promoting internet application. JMIR Res Protoc 2016 Mar 04;5(1):e36. [doi: 10.2196/resprot.5232]
9. Marin-Torres V, Valverde Aliaga J, Sánchez Miró I, Sáenz Del Castillo Vicente MI, Polentinos-Castro E, Garrido Barral A. [Internet as an information source for health in primary care patients and its influence on the physician-patient relationship]. Aten Primaria 2013 Jan;45(1):46-53 [FREE Full text] [doi: 10.1016/j.aprim.2012.09.004] [Medline: 23140836]
10. Savage SV, Kwan S, Bergstrand K. Virtual health: the impact of health-related websites on doctor-patient interactions. In: Kronenfeld JJ, editor. Technology, Communication, Disparities and Government Options in Health and Health Care Services. Bingley, UK: Emerald Group Publishing Limited; 2014:93-116.
11. Laugesen J, Hassanein K, Yuan Y. The impact of internet health information on patient compliance: a research model and an empirical study. J Med Internet Res 2015 Jun 11;17(6):e143 [FREE Full text] [doi: 10.2196/jmir.4333] [Medline: 26068214]
12. Roper L, Jorm C. Please leave your phone outside: policymakers and medical app providers are encouraging patient participation, but doctors in the consult room are not. BMJ Innov 2017 Mar 24;3(2):65-70. [doi: 10.1136/bmjinnov-2016-000134]
13. Weiner JP. Doctor-patient communication in the e-health era. Isr J Health Policy Res 2012;1(1):33. [doi: 10.1186/2045-4015-1-33]
14. 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 Jun 22;17(6):e155 [FREE Full text] [doi: 10.2196/jmir.3792] [Medline: 26099267]
15. Riessman RC. Narrative Analysis (Qualitative Research Methods Series 30). Newbury Park, CA: Sage Publications; 1993.

16. Charmaz K. Grounded theory: objectivist constructivist methods. In: Denzin NK, Lincoln YS, editors. Handbook of Qualitative Research. Thousand Oaks, CA: SAGE Publications, Inc; 2000:509-535.
17. Moore J, Engel J, Prentice D. Relational ethics in everyday practice. Can Oncol Nurs J 2014;24(1):31-39. [Medline: 24707705]
18. Riessman CK. Narrative Analysis. In: Lewis-Beck MS, Bryman A, Liao TF, editors. The SAGE Encyclopedia of Social Science Research Methods. Thousand Oaks, CA: Sage Publications Inc; 2003:706-709.
19. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
20. Coffey AJ, Atkinson PA. Narratives and stories. In: Coffey AJ, Atkinson PA, editors. Making Sense of Qualitative Data: Complimentary Research Strategies (And Social Thought). Thousand Oaks, CA: Sage publications; 1996:54-82.
21. Bergum V, Dossetor J. Relational Ethics: The Full Meaning of Respect. Hagerstown, MD: University Publishing Group; 2005.
22. Lupton D. M-health and health promotion: the digital cyborg and surveillance society. Soc Theory Health 2012 Jun 27;10(3):229-244. [doi: 10.1057/sth.2012.6]
23. Ocloo J, Matthews R. From tokenism to empowerment: progressing patient and public involvement in healthcare improvement. BMJ Qual Saf 2016 Aug;25(8):626-632 [FREE Full text] [doi: 10.1136/bmjqs-2015-004839] [Medline: 26993640]
24. Zolnierek KB, Dimatteo MR. Physician communication and patient adherence to treatment: a meta-analysis. Med Care 2009 Aug;47(8):826-834 [FREE Full text] [doi: 10.1097/MLR.0b013e31819a5acc] [Medline: 19584762]
25. Aujoulat I, Marcolongo R, Bonadiman L, Deccache A. Reconsidering patient empowerment in chronic illness: a critique of models of self-efficacy and bodily control. Soc Sci Med 2008 Mar;66(5):1228-1239. [doi: 10.1016/j.socscimed.2007.11.034] [Medline: 18155338]
26. 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]
27. Austin W. Moral distress and the contemporary plight of health professionals. HEC Forum 2012 Mar;24(1):27-38. [doi: 10.1007/s10730-012-9179-8] [Medline: 22441996]

Abbreviations

HCPs: health care professionals

Edited by G Eysenbach; submitted 16.09.17; peer-reviewed by L Roper, A Syrowatka; comments to author 08.11.17; revised version received 17.11.17; accepted 17.11.17; published 26.01.18.

Please cite as:
Macdonald GG, Townsend AF, Adam P, Li LC, Kerr S, McDonald M, Backman CL
eHealth Technologies, Multimorbidity, and the Office Visit: Qualitative Interview Study on the Perspectives of Physicians and Nurses
J Med Internet Res 2018;20(1):e31
URL: http://www.jmir.org/2018/1/e31/
doi:10.2196/jmir.8983
PMID:29374004

Review

Evaluation Criteria of Noninvasive Telemonitoring for Patients With Heart Failure: Systematic Review

Troskah Farnia[1], MSc; Marie-Christine Jaulent[1], PhD; Olivier Steichen[1,2], MD, PhD

[1]Laboratoire d'Informatique Médicale et Ingénierie des Connaissances en eSanté, Institut National de la Santé et de la Recherche Médicale, Sorbonne Universités, Université Paris 13, Sorbonne Paris Cité, Paris, France

[2]Department of Internal Medicine, Hôpital Tenon, Assistance Publique-Hôpitaux de Paris, Paris, France

Corresponding Author:
Olivier Steichen, MD, PhD
Department of Internal Medicine
Hôpital Tenon
Assistance Publique-Hôpitaux de Paris
4 Rue de la Chine
Paris,
France
Phone: 33 1 56 01 78 31
Email: olivier.steichen@aphp.fr

Abstract

Background: Telemonitoring can improve heart failure (HF) management, but there is no standardized evaluation framework to comprehensively evaluate its impact.

Objective: Our objectives were to list the criteria used in published evaluations of noninvasive HF telemonitoring projects, describe how they are used in the evaluation studies, and organize them into a consistent scheme.

Methods: Articles published from January 1990 to August 2015 were obtained through MEDLINE, Web of Science, and EMBASE. Articles were eligible if they were original reports of a noninvasive HF telemonitoring evaluation study in the English language. Studies of implantable telemonitoring devices were excluded. Each selected article was screened to extract the description of the telemonitoring project and the evaluation process and criteria. A qualitative synthesis was performed.

Results: We identified and reviewed 128 articles leading to 52 evaluation criteria classified into 6 dimensions: clinical, economic, user perspective, educational, organizational, and technical. The clinical and economic impacts were evaluated in more than 70% of studies, whereas the educational, organizational, and technical impacts were studied in fewer than 15%. User perspective was the most frequently covered dimension in the development phase of telemonitoring projects, whereas clinical and economic impacts were the focus of later phases.

Conclusions: Telemonitoring evaluation frameworks should cover all 6 dimensions appropriately distributed along the telemonitoring project lifecycle. Our next goal is to build such a comprehensive evaluation framework for telemonitoring and test it on an ongoing noninvasive HF telemonitoring project.

(J Med Internet Res 2018;20(1):e16) doi:10.2196/jmir.7873

KEYWORDS

telemedicine; outcome and process assessment (health care); program evaluation; heart failure

Introduction

Heart failure (HF) affects 26 million people worldwide, incurring direct and indirect costs of more than US $100 billion per year [1,2]. HF causes 1% to 2% of all hospitalizations, representing a major burden for patients and the health care system [3]. Number of hospitalizations is a marker of disease instability, and studies have shown that HF management can reduce this number [4].

Telemonitoring is a branch of telemedicine defined as the use of communication technologies to monitor and transmit data on the health status of patients to distant care providers [5,6]. It differs from teleconsultation, where there is a real-time interactive video or audio consultation between the patient and a distant health care provider. It also differs from tele-expertise, where a health care provider presents a patient case and gets advice from a distant colleague through a dedicated system, without direct patient involvement, in real or deferred time.

Systematic reviews of telemonitoring projects have shown reduced hospitalization rates for acute HF [7-9]. Both the European Society of Cardiology and the American Heart Association support the use of telemonitoring to improve the care of HF patients [10,11]. HF telemonitoring can involve the use of invasive or noninvasive monitoring devices. Invasive devices are implanted in the body, and data transmission is not controlled by the patient. Noninvasive devices, like weight scales or blood pressure monitors, are used, mostly by the patient, to self-monitor physiological measurements, signs, or symptoms of the disease.

The information flow is much more intricate with noninvasive than with invasive telemonitoring. First, human action is needed to handle noninvasive telemonitoring devices whereas monitored data is automatically recorded and sent by invasive telemonitoring devices. Second, patients (or family caregivers) are major actors of the noninvasive telemonitoring process whereas their involvement is minimal with invasive telemonitoring. Third, nurses and primary care physicians are usually involved in the noninvasive telemonitoring process whereas they do not take part in invasive monitoring. Fourth, as a result of these differences, noninvasive HF telemonitoring requires readiness for change, education, and training of patients and caregivers whereas invasive HF telemonitoring does not [8]. Due to these major differences, this review focuses on noninvasive HF telemonitoring.

The development of evaluation criteria for electronic tools is considered to be a critical step by the European Society of Cardiology [10]. Standard evaluation frameworks are useful to encourage systematic evaluation and get conclusive results that can be compared or aggregated across programs, allowing the analysis of determinants of success and failure for efficient resource allocation. Standard evaluation frameworks have been used for the evaluation of teleconsultation and tele-expertise [12-16]. To our knowledge, they have not yet been used for the evaluation of telemonitoring projects. Telemonitoring interventions are complex: they involve many different actors with different backgrounds (health care professional, patients and family, technicians, payers), they use technical devices, and they change the usual process of care. They can impact health care on many levels [17]: patient access to care, health and quality of life, patient and care provider education, family and care provider workload, organization of the patient care pathway, health care costs, and more. A comprehensive telemonitoring evaluation framework therefore needs to be multidimensional.

Our aim was to perform a systematic review of criteria used for the evaluation of noninvasive HF telemonitoring projects, describe how they are used in evaluation studies, and organize them into a consistent scheme.

Methods

Information Sources and Eligibility Criteria

We did not submit a review protocol to a prospective register. We searched Medical Literature Analysis and Retrieval System Online (MEDLINE), Excerpta Medica database (EMBASE), and Web of Science for articles published from January, 1990, to the query date (August 15, 2015) using the queries displayed in Multimedia Appendix 1. The search strategy for each bibliographic database was internally discussed, piloted, and refined by the authors but not submitted for external peer review. We checked reference lists of included articles to identify additional studies. We also took advantage of 3 Cochrane reviews on telemonitoring, published shortly after the last query date of our review, to look for missed studies [18-20].

Articles were eligible if they were original reports of a noninvasive telemonitoring project evaluation study for HF using explicit evaluation criteria. Articles were excluded if they did not contain original data (reviews, editorials, position papers, etc), were not written in English, focused on other types of telemedicine (teleconsultation, tele-expertise, etc), and if they reported invasive telemonitoring for HF through implantable devices.

Study Selection and Data Collection

After eliminating duplicate articles, titles and abstracts were independently screened by 2 readers to exclude obviously irrelevant articles. Discordant classifications between the 2 readers were resolved through discussion. The full text of remaining articles was read by 1 investigator, who applied eligibility and exclusion criteria. The final selection was cross-checked by a second investigator. Characteristics of the telemonitoring project, characteristics of the evaluation process, and evaluation criteria were systematically abstracted by 1 investigator and cross-checked by another. The collected data are reported in Multimedia Appendix 2.

Synthesis of Results

A preliminary list of broad evaluation dimensions was adapted by 2 investigators from previously published evaluation frameworks for telemedicine [12-16]. This categorization was then iteratively refined to meld the evaluation criteria found in each reviewed study into a consistent scheme.

Results

Study Selection and Characteristics

The queries of bibliographic databases identified 328 potentially eligible articles, and we included 128 articles in the review (reference list is in Multimedia Appendix 3). The review flowchart describes the process and reasons for exclusion (Figure 1). Characteristics of the 128 studies are reported in Multimedia Appendix 4, and a summary is presented in Table 1.

Europe and the United States contributed the most to the assessment of HF telemonitoring (50% of studies performed in Europe and 41% in the United States). The first study began in

1997 in the United States [21]. Europe started to carry out research to assess HF telemonitoring 3 years later [22-24].

Telemonitoring Characteristics

Some features were highly prevalent across the telemonitoring projects: 80% were carried out by a cardiology team, and the care providers were most often HF nurses (86% of projects). Patients were actively involved in 100% of the projects, but the family of the patients, psychologists, and technicians rarely participated. In 75% of studies, patients were included in the telemonitoring program at discharge from a hospitalization for acute heart failure. The phase of the project lifecycle was clear and identifiable in 125 articles: most of these projects were in the implementation phase (59%), and no project was part of routine clinical care.

The primary monitored data in the telemonitoring projects were weight, HF symptoms, heart rate, and blood pressure. These data were transmitted via telephone (verbal communication or keypad) or Internet (mobile phone or tablet). If the monitored data fell outside predefined boundaries, a warning was triggered and led to corrective actions. However, these actions were described in only 63 articles (49% of studies).

Evaluation Dimensions and Criteria

We retrieved 52 criteria from the 128 studies (Textbox 1) and classified them into 6 main dimensions: economic, clinical, educational, technical, user perspective, and organizational.

Most studies (95%) covered, at most, 3 dimensions, and none covered all 6 (Multimedia Appendix 5). Clinical and economic dimensions were assessed in over 70% of studies, whereas the educational, organizational, and technical dimensions were studied in less than 15% (Table 2).

The evaluation dimensions were not used homogeneously across all phases of project lifecycle. User perceptive was the most often covered dimension in the development phase with clinical and economic dimensions covered most in the later phases (implementation and integration).

Each dimension includes from 2 to 16 criteria (Textbox 1). The total number of criteria used per study ranged from 1 to 11 (Multimedia Appendix 6). The most often used evaluation criteria were cost and resource utilization (71%) and quality of life (51%) (Table 3). The criteria within the same dimension were also not used homogeneously across all phases of the project lifecycle. Concerning user perspective, for example, ease of use of the devices was more often evaluated in the development phase, whereas satisfaction with care was more often evaluated in the later phases of the project lifecycle (Table 3).

Figure 1. Review flowchart.

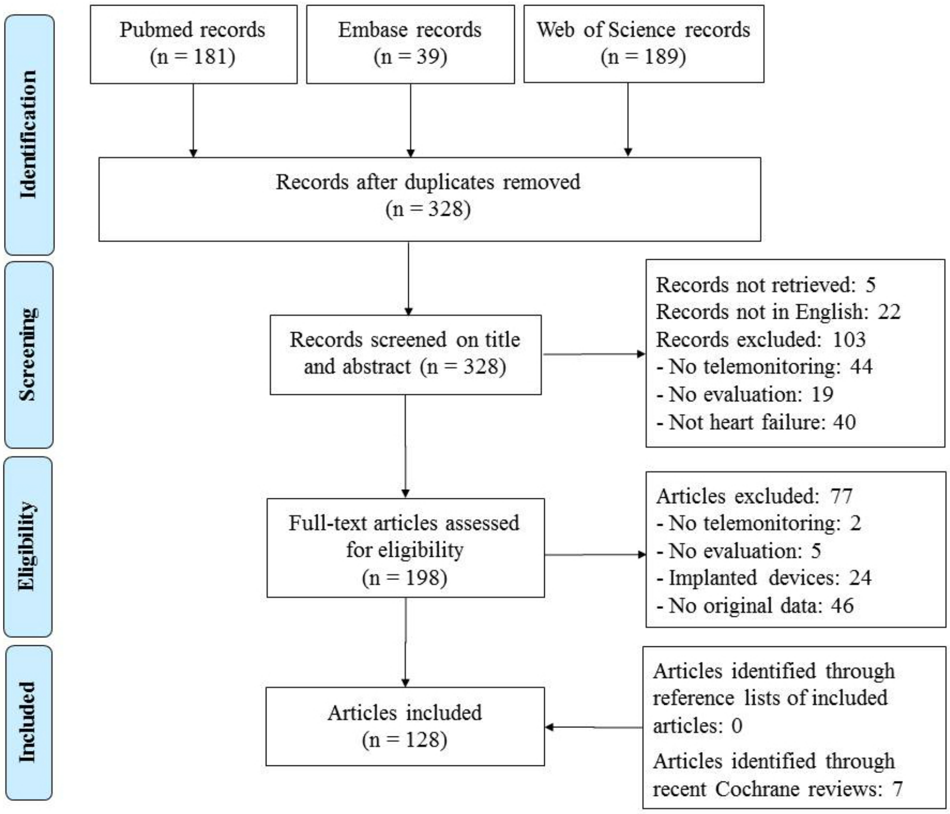


Table 1. Study characteristics.

Variable	Description	Value
Start of the experiment	Begins with the recruitment of the first patient.	Median: 2006 (IQR[a] 1999, 2004) Unclear: 40% (51 studies)
Country	Country where the experiment was carried out.	Europe: 50% (64 studies) United States: 41% (53 studies) Canada: 4% (5 studies) Australia: 3% (4 studies) Asia: 1.5% (2 studies)
Care context	Location of the patients within their care pathway (after a hospitalization for an acute episode or in stable condition, followed-up by a specialist or a primary care provider).	Post-acute hospitalization: 75% (96 studies) Cardiology team: 87% (111 studies)
Telemonitoring actors	People directly involved in the telemonitoring service process (patients, health care professionals, nurses, care givers, technicians).	Patients: 100% (128 studies) Nurses: 86% (110 studies) Cardiologists: 46% (59 studies) General physicians: 26% (33 studies) Psychologists: 4% (5 studies) Families: 4% (5 studies) Technicians: 2.3% (3 studies)
Project lifecycle phase	Four phases depict the lifecycle of a project: technical development (prototype), clinical implementation (small scale experiment), integration in the clinical pathway (large scale experiment), and routine patient care.	Implementation: 59% (76 studies) Integration: 20% (26 studies) Development: 18% (23 studies) Routine care: 0% (no study) Unclear: 2.3% (3 studies)
Study design	Methodological design of the study (randomized trial, cohort study, cross-sectional study, etc).	Randomized controlled trial: 62% (79 studies) Cohort: 34% (43 studies) Cross-sectional: 3% (4 studies) Unclear: 1.5% (2 studies)
Patient age	Mean age of included patients.	Median: 69 (IQR 62.8, 74.2) years Unclear: 18% (23 studies)
Number of patients	Total number of subjects involved in the telemonitoring assessment.	Patients: median 157.5 (IQR 71, 376) Volunteers: median 17 (IQR 11, 24) Unclear: 6% (7 studies)
Number of health care professionals	Total number of health care professionals involved in the telemonitoring assessment.	Care partners: median 52 (IQR 32, 82) Nurses: median 4 (IQR 3,4) Cardiologists: median 13 (IQR 6, 20) Primary care physicians: median 260 (IQR 181, 313) Unclear: 92% (118 studies)
Associated interventions	Interventions performed to enable the home telemonitoring service (therapeutic education, training in equipment use, etc).	Therapeutic education: 32% (41 studies) Training of use of equipment: 30% (38 studies) Informed family: 4% (5 studies) Home nurse visit: 2.3% (3 studies) Technical support: 0.7% (1 study) Unclear: 46% (59 studies)
Intervention duration	Duration of the monitoring service per patient, from the first to the last data transmission.	Median: 9 (IQR 6,12) months Unclear: 10% (13 studies)

[a] IQR: interquartile range.

Textbox 1. Evaluation dimensions and related criteria.

Economic

- Costs of resource utilization: days in the hospital, number of nurse visits, number of consultations with cardiologist and general practitioner, number of emergency visits, hospitalization, and readmission rate
- Cost of technical development, deployment, and maintenance of the service

Clinical

- Patient-reported outcomes: quality of life, health status, functional capacity, and activities of daily living
- Disease-oriented outcomes: mortality rate and morbidity rate
- Patient adherence to treatment: diet and medication
- Physician-adherence to guidelines

Educational

- Patient knowledge of the disease
- Patient self-care knowledge and behavior
- Primary care physician education
- Familial caregiver involvement

Technical

- Ergonomics: intuitive functions and design, quick on/off switch, and setup and configuration of the system
- Characteristics: platform connection with other devices, authentication, secure storage, maintainability, and availability of service

User perspective

- Patient perception: feelings of patient, cognitive feedback, acceptability of technology and service, reliability of information and communication technology, willingness to pay, patient motivation, social network, self-efficacy and confidence, adaptation to telephone monitoring, ease of use, access to care providers, satisfaction with new technology, compliance with new technology, and overall satisfaction with the process of care
- Care provider perception: satisfaction, utility, acceptability of technology, ease of use, compliance with new technology, and overall satisfaction with the process of care

Organizational

- Administrative: insurance policy and hospital policy
- Clinical: acceptability of heart failure nurses by general physician, heart failure nurse/ physician communication, patient/physician communication, and physician workload

Table 2. Coverage of evaluation dimensions across studies and phases of the project lifecycle.

Dimension	Number of studies, n (%)	Lifecycle phase		
		Development (n=23), n (%)	Implementation (n=76), n (%)	Integration (n=26), n (%)
Clinical	107 (84)	11 (48)	68 (89)	23 (88)
Economic	91 (71)	9 (39)	57 (75)	22 (85)
User perspective	55 (43)	19 (83)	31 (41)	12 (46)
Educational	18 (14)	3 (13)	10 (13)	5 (19)
Organizational	7 (5)	1 (4)	5 (7)	1 (4)
Technical	6 (4)	3 (13)	3 (4)	0 (0)

Table 3. Evaluation criterion most frequently used in each dimension.

Criterion	Dimension	Overall (n=128), n (%)	Lifecycle phase		
			Development (n=23), n (%)	Implementation (n=76), n (%)	Integration (n=26), n (%)
Cost of resources utilization	Economic	91 (71)	9 (39)	58 (76)	22 (85)
Quality of life	Clinical outcomes	65 (51)	8 (35)	42 (55)	14 (54)
Patient and family satisfaction with new technology	User perspective	21 (16)	7 (30)	10 (13)	3 (12)
Knowledge of disease	Educational outcomes	14 (11)	3 (13)	8 (10.5)	4 (15)
Patient and physician communication	Organizational	3 (2.3)	1 (4)	1 (1.3)	1 (4)
Reliability of transmitted data	Technical	2 (1.6)	2 (9)	0 (0)	0 (0)
Device specifications	Technical	2 (1.6)	2 (9)	0 (0)	0 (0)

Discussion

We found 128 studies using a total of 52 evaluation criteria categorized into 6 high-level dimensions. No study covered all 6 evaluation dimensions. The evaluation dimensions were not used with the same frequency for all phases of the project lifecycle. The principle focus in the development phase was on user perspective, whereas the focus in the latter phases of the lifecycle was on the clinical and economical dimensions. The technical, organizational, and educational dimensions were poorly evaluated overall.

The 6 dimensions were derived from telemedicine assessment frameworks *Grille d'Evaluation Multidisciplinaire Santé Autonomie* (multidimensional evaluation grid for health and autonomy) [12], model for assessment of telemedicine [13], Khoja-Durrani-Scott evaluation framework [14], *Technologique, Ergonomique, Médicale, Sociale, Économique et Déontologique* (technological, ergonomic, medical, social, economic, and ethical) [15], and the 3-dimensional model [16]. These frameworks required adaptations to better fit telemonitoring. First, the role of patients and nurses is prominent in telemonitoring projects, whereas it is more limited in teleconsultation and tele-expertise. Thus, smooth collaboration must be ensured and evaluated between patients and care providers as much as between care providers themselves. Specific organizational and educational evaluation criteria are therefore needed for telemonitoring. Second, telemonitoring stands out from a technical point of view because devices are needed to gather data. The ergonomic assessment, user perception, and technical characteristics of these devices are thus key elements in the assessment of telemonitoring projects [15].

Previous evaluation frameworks are not adapted to all phases of a telemonitoring project lifecycle and do not take into account the perspective of all telemonitoring actors and external stakeholders (manufacturers, payers, etc). For example, a telemonitoring project at the first phase of its lifecycle (pilot) cannot appropriately evaluate clinical outcomes and, at the other end, proper technical evaluation is a prerequisite long before the last phase of a telemonitoring project lifecycle (routine clinical care). Criteria and indicators for a given dimension will differ across lifecycle phases. For example, the user perspective should be evaluated in all 4 phases but with different criteria: “ease of use of the system” is an appropriate criterion in the development phase, “satisfaction with new technology” in the implementation and integration phases, and “overall satisfaction with the process of care” in routine clinical care.

This review is limited by its focus on noninvasive HF telemonitoring. However, aside from disease-specific clinical outcomes, noninvasive telemonitoring services share many technical, economic, organizational, and educational features independently of the target disease. Our framework should therefore be easy to adapt to other health conditions. Our search strategy may have missed some studies that were described by other keywords. However, we found only 7 additional studies in 3 recent Cochrane reviews [18-20]. The number of missing studies is therefore likely to be low. Our 52 criteria cover all outcomes reported in previous systematic reviews on the evaluation of telemonitoring in HF [18,25] but also in other chronic diseases, such as diabetes and chronic obstructive pulmonary diseases [19,20]. We have identified broad evaluation dimensions from previously published evaluation frameworks for telemedicine projects and refined these dimensions iteratively during the review process. The 6 final dimensions accommodate all evaluation criteria used in previously published evaluation studies, but other categorization schemes are certainly possible and should be compared. The data and methods used in this review were not suited to assess the strength, limitations, relevance, and usefulness of each retrieved evaluation criterion. A follow-up to our work is needed to provide more guidance for the use of criteria in future HF telemonitoring evaluation studies.

Comprehensive telemonitoring evaluation frameworks should cover all 6 dimensions and help users choose the appropriate dimensions and evaluation criteria depending on the phase of their telemonitoring project lifecycle and the perspective of telemonitoring actors or external stakeholders they want to adopt. Our next goal is to build such a framework for noninvasive HF telemonitoring, deliberately emphasizing the technical, organizational, and educational dimensions that have been neglected by previous telemonitoring assessment studies. We will test this framework on an ongoing HF telemonitoring project.

Acknowledgments

This work was supported by the Regional Health Agency of Centre Val de Loire and the French National Association of Technical Research (ANRT).

Conflicts of Interest

None to declare.

Multimedia Appendix 1

Bibliographic search strategy.

[PDF File (Adobe PDF File), 16KB - jmir_v20i1e16_app1.pdf]

Multimedia Appendix 2

Collected study characteristics.

[PDF File (Adobe PDF File), 18KB - jmir_v20i1e16_app2.pdf]

Multimedia Appendix 3

Included papers reference list.

[PDF File (Adobe PDF File), 965KB - jmir_v20i1e16_app3.pdf]

Multimedia Appendix 4

Detailed study characteristics.

[XLS File (Microsoft Excel File), 181KB - jmir_v20i1e16_app4.xls]

Multimedia Appendix 5

Number of evaluation dimensions covered per study.

[PPT File (Microsoft PowerPoint Presentation), 124KB - jmir_v20i1e16_app5.ppt]

Multimedia Appendix 6

Number of evaluation criteria used per study.

[PPT File (Microsoft PowerPoint Presentation), 138KB - jmir_v20i1e16_app6.ppt]

Multimedia Appendix 7

Collected study characteristics.

[PDF File (Adobe PDF File), 18KB -]

References

1. Cook C, Cole G, Asaria P, Jabbour R, Francis DP. The annual global economic burden of heart failure. Int J Cardiol 2014 Feb 15;171(3):368-376. [doi: 10.1016/j.ijcard.2013.12.028] [Medline: 24398230]
2. Ambrosy AP, Fonarow GC, Butler J, Chioncel O, Greene SJ, Vaduganathan M, et al. The global health and economic burden of hospitalizations for heart failure: lessons learned from hospitalized heart failure registries. J Am Coll Cardiol 2014 Apr 01;63(12):1123-1133 [FREE Full text] [doi: 10.1016/j.jacc.2013.11.053] [Medline: 24491689]
3. Berry C, Murdoch D, McMurray J. Economics of chronic heart failure. Eur J Heart Fail 2001 Jun;3(3):283-291 [FREE Full text] [Medline: 11377998]
4. Smith B, Forkner E, Zaslow B, Krasuski RA, Stajduhar K, Kwan M, et al. Disease management produces limited quality-of-life improvements in patients with congestive heart failure: evidence from a randomized trial in community-dwelling patients. Am J Manag Care 2005 Nov;11(11):701-713 [FREE Full text] [Medline: 16268753]
5. Louis A, Turner T, Gretton M, Baksh A, Cleland J. A systematic review of telemonitoring for the management of heart failure. Eur J Heart Fail 2003 Oct;5(5):583-590 [FREE Full text] [Medline: 14607195]

6. Meystre S. The current state of telemonitoring: a comment on the literature. Telemed J E Health 2005 Feb;11(1):63-69. [doi: 10.1089/tmj.2005.11.63] [Medline: 15785222]
7. Chaudhry SI, Phillips CO, Stewart SS, Riegel B, Mattera JA, Jerant AF, et al. Telemonitoring for patients with chronic heart failure: a systematic review. J Card Fail 2007 Feb;13(1):56-62 [FREE Full text] [doi: 10.1016/j.cardfail.2006.09.001] [Medline: 17339004]
8. Ramaekers B, Janssen-Boyne J, Gorgels A, Vrijhoef H. Adherence among telemonitored patients with heart failure to pharmacological and nonpharmacological recommendations. Telemed J E Health 2009;15(6):517-524. [doi: 10.1089/tmj.2009.0160] [Medline: 19566401]
9. Kotb A, Cameron C, Hsieh S, Wells G. Comparative effectiveness of different forms of telemedicine for individuals with heart failure (HF): a systematic review and network meta-analysis. PLoS One 2015;10(2):e0118681 [FREE Full text] [doi: 10.1371/journal.pone.0118681] [Medline: 25714962]
10. Cowie M, Bax J, Bruining N, Cleland J, Koehler F, Malik M, et al. e-Health: a position statement of the European Society of Cardiology. Eur Heart J 2016 Jan 01;37(1):63-66. [doi: 10.1093/eurheartj/ehv416] [Medline: 26303835]
11. Schwamm L, Chumbler N, Brown E, Fonarow G, Berube D, Nystrom K, American Heart Association Advocacy Coordinating Committee. Recommendations for the Implementation of Telehealth in Cardiovascular and Stroke Care: A Policy Statement From the American Heart Association. Circulation 2017 Feb 14;135(7):e24-e44 [FREE Full text] [doi: 10.1161/CIR.0000000000000475] [Medline: 27998940]
12. Le Goff-Pronost M, Picard R. Need for ICTs assessment in the health sector: a multidimensional framework. Commun Strateg 2011;83(3):87-108 [FREE Full text]
13. Kidholm K, Ekeland A, Jensen L, Rasmussen J, Pedersen C, Bowes A, et al. A model for assessment of telemedicine applications: MAST. Int J Technol Assess Health Care 2012 Jan;28(1):44-51. [doi: 10.1017/S0266462311000638] [Medline: 22617736]
14. Khoja S, Durrani H, Scott R, Sajwani A, Piryani U. Conceptual framework for development of comprehensive e-health evaluation tool. Telemed J E Health 2013 Jan;19(1):48-53. [doi: 10.1089/tmj.2012.0073] [Medline: 22957502]
15. Fouquet Y, Sablier J, Vuillerme N, Daynes P, Rialle V. TEMSED: une approche multicritère d'évaluation des technologies de télésanté et d'autonomie. Ann Phys Rehabil Med 2011 Oct;54(supplement 1):e57. [doi: 10.1016/j.rehab.2011.07.776]
16. Bashshur R, Shannon G, Sapci H. Telemedicine evaluation. Telemed J E Health 2005 Jun;11(3):296-316. [doi: 10.1089/tmj.2005.11.296] [Medline: 16035928]
17. Gurné O, Conraads V, Missault L, Mullens W, Vachierys J, Van Meighem W, et al. A critical review on telemonitoring in heart failure. Acta Cardiol 2012 Aug;67(4):439-444. [Medline: 22997998]
18. Inglis S, Clark R, Dierckx R, Prieto-Merino D, Cleland J. Structured telephone support or non-invasive telemonitoring for patients with heart failure. Cochrane Database Syst Rev 2015;10:CD007228. [doi: 10.1002/14651858]
19. Posadzki P, Mastellos N, Ryan R, Gunn LH, Felix LM, Pappas Y, et al. Automated telephone communication systems for preventive healthcare and management of long-term conditions. Cochrane Database Syst Rev 2016 Dec 14;12:CD009921. [doi: 10.1002/14651858.CD009921.pub2] [Medline: 27960229]
20. 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]
21. Benatar D, Bondmass M, Ghitelman J, Avitall B. Outcomes of Chronic Heart Failure. Arch Intern Med 2003 Feb 10;163(3):347-352. [doi: 10.1001/archinte.163.3.347] [Medline: 12578516]
22. Ferrante D, Varini S, Macchia A, Soifer S, Badra R, Nul D, et al. Long-term results after a telephone intervention in chronic heart failure: DIAL (Randomized Trial of Phone Intervention in Chronic Heart Failure) follow-up. J Am Coll Cardiol 2010 Jul 27;56(5):372-378 [FREE Full text] [doi: 10.1016/j.jacc.2010.03.049] [Medline: 20650358]
23. Giordano A, Zanelli E, Scalvini S. Home-based telemanagement in chronic heart failure: an 8-year single-site experience. J Telemed Telecare 2011;17(7):382-386. [doi: 10.1258/jtt.2011.110201] [Medline: 21979603]
24. Cleland JGF, Louis AA, Rigby AS, Janssens U, Balk AHMM. Noninvasive home telemonitoring for patients with heart failure at high risk of recurrent admission and death: the Trans-European Network-Home-Care Management System (TEN-HMS) study. J Am Coll Cardiol 2005 May 17;45(10):1654-1664 [FREE Full text] [doi: 10.1016/j.jacc.2005.01.050] [Medline: 15893183]
25. Kitsiou S, Paré G, Jaana M. Effects of home telemonitoring interventions on patients with chronic heart failure: an overview of systematic reviews. J Med Internet Res 2015 Mar;17(3):e63 [FREE Full text] [doi: 10.2196/jmir.4174] [Medline: 25768664]

Abbreviations

EMBASE: Excerpta Medica database
HF: heart failure
MEDLINE: Medical Literature Analysis and Retrieval System Online

Edited by G Eysenbach; submitted 26.04.17; peer-reviewed by S Agboola, JB Park; comments to author 12.07.17; revised version received 18.10.17; accepted 20.11.17; published 16.01.18.

Please cite as:
Farnia T, Jaulent MC, Steichen O
Evaluation Criteria of Noninvasive Telemonitoring for Patients With Heart Failure: Systematic Review
J Med Internet Res 2018;20(1):e16
URL: http://www.jmir.org/2018/1/e16/
doi:10.2196/jmir.7873
PMID:29339348

Original Paper

Effect of Contract Compliance Rate to a Fourth-Generation Telehealth Program on the Risk of Hospitalization in Patients With Chronic Kidney Disease: Retrospective Cohort Study

Chi-Sheng Hung[1], MD, PhD; Jenkuang Lee[1], MD, PhD; Ying-Hsien Chen[1], MD; Ching-Chang Huang[1], MD; Vin-Cent Wu[2], MD, PhD; Hui-Wen Wu[1], RN; Pao-Yu Chuang[1]; Yi-Lwun Ho[1], MD, PhD

[1]Telehealth Center, National Taiwan University Hospital, Taipei, Taiwan

[2]Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan

Corresponding Author:
Yi-Lwun Ho, MD, PhD
Telehealth Center, National Taiwan University Hospital
#7, Chung-Shan South Road
Taipei,
Taiwan
Phone: 886 2 2312 3456
Fax: 886 2 27044688
Email: ylho@ntu.edu.tw

Abstract

Background: Chronic kidney disease (CKD) is prevalent in Taiwan and it is associated with high all-cause mortality. We have shown in a previous paper that a fourth-generation telehealth program is associated with lower all-cause mortality compared to usual care with a hazard ratio of 0.866 (95% CI 0.837-0.896).

Objective: This study aimed to evaluate the effect of renal function status on hospitalization among patients receiving this program and to evaluate the relationship between contract compliance rate to the program and risk of hospitalization in patients with CKD.

Methods: We retrospectively analyzed 715 patients receiving the telehealth care program. Contract compliance rate was defined as the percentage of days covered by the telehealth service before hospitalization. Patients were stratified into three groups according to renal function status: (1) normal renal function, (2) CKD, or (3) end-stage renal disease (ESRD) and on maintenance dialysis. The outcome measurements were first cardiovascular and all-cause hospitalizations. The association between contract compliance rate, renal function status, and hospitalization risk was analyzed with a Cox proportional hazards model with time-dependent covariates.

Results: The median follow-up duration was 694 days (IQR 338-1163). Contract compliance rate had a triphasic relationship with cardiovascular and all-cause hospitalizations. Patients with low or very high contract compliance rates were associated with a higher risk of hospitalization. Patients with CKD or ESRD were also associated with a higher risk of hospitalization. Moreover, we observed a significant interaction between the effects of renal function status and contract compliance rate on the risk of hospitalization: patients with ESRD, who were on dialysis, had an increased risk of hospitalization at a lower contract compliance rate, compared with patients with normal renal function or CKD.

Conclusions: Our study showed that there was a triphasic relationship between contract compliance rate to the telehealth program and risk of hospitalization. Renal function status was associated with risk of hospitalization among these patients, and there was a significant interaction with contract compliance rate.

(J Med Internet Res 2018;20(1):e23) doi:10.2196/jmir.8914

KEYWORDS

telehealth; contract compliance rate; chronic kidney disease

Introduction

Telehealth has been increasingly used in the management of chronic conditions such as diabetes, asthma, and cardiovascular disease (CVD) [1-4]. A fourth-generation telehealth program is an Internet-based, synchronized disease management program using telemonitoring technology to provide an immediate response. In a prior report, we have shown that 576 patients with chronic CVD who received a fourth-generation telehealth care program had lower mortality compared to 1178 patients who received usual care [5]. The adjusted hazard ratio of all-cause mortality for the use of telehealth was 0.866 (95% CI 0.837-0.896). Based on these results, this technology has been increasingly applied to the care of other chronic conditions in Taiwan.

Chronic kidney disease (CKD) is a chronic condition characterized by a decreased glomerular filtration rate (GFR; <60 mL/min/1.73 m^2) and associated with risks of progressive deterioration in kidney function and of CVD [6]. To halt the deterioration in renal function, the Kidney Disease: Improving Global Outcomes organization published CKD management guidelines, which recommend treatment of the causes of the disease and of factors that may aggravate kidney deterioration [7]. Compliance to the guidelines and recommended treatments is crucial for the long-term management of patients with CKD [8]. Telehealth programs have potential advantages in ensuring treatment compliance and early recognition of disease complications. Telehealth initiatives in which nephrology specialists provide their expertise remotely over the Internet represent a model that can be adopted in low-resource settings [9]. There have been a few studies on the use of telemedicine among patients receiving dialysis [10,11] or renal transplant recipients [12], but none on the use of telehealth programs for the management of patients with stage 3 to 4 CKD.

Chronic kidney disease is prevalent in Taiwan and is associated with high all-cause mortality [13]. We have shown that a fourth-generation telehealth program is associated with lower mortality compared with usual care [5]. However, the influence of renal function status on hospitalization among patients receiving the program has not previously been evaluated. Similarly, the relationship between contract compliance rate to the program and risk of hospitalization in patients with CKD has not yet been reported. The aim of this study was to elucidate if the contract compliance rate to the telehealth program is associated with a higher rate of hospitalization among patients with or without CKD, and to identify any possible interactions between contract compliance rate and CKD. We designed this retrospective cohort study to answer this question.

Methods

Study Design

This was a single-center, retrospective study that was approved by the Institutional Review Board of National Taiwan University Hospital, Taipei, Taiwan. Informed consent was obtained from all participants.

Recruitment

The study was conducted from December 2009 to April 2013 at the Telehealth Center of the hospital by the Taiwan ELEctroHEALTH Study Group (TELEHEALTH Study Group). Patients older than 20 years with diagnosed chronic CVD and receiving the telehealth program at our telehealth center were enrolled as the study group. The decision of whether to receive the telehealth program depended on the patients and/or their caregivers. Chronic CVD included coronary artery disease, myocardial infarction, heart failure, peripheral artery disease, stroke, and hypertension.

Definition of Chronic Kidney Disease

The definition of CKD is based on the Kidney Disease Outcomes Quality Initiative guidelines published by the National Kidney Foundation [14]. Chronic kidney disease is defined as a GFR of less than 60 mL/min/1.73 m^2. The renal function status is classified into (1) normal: no kidney damage, with a GFR ≥60 mL/min/1.73 m^2; (2) CKD: kidney damage, with GFR <60 mL/min/1.73 m^2; or (3) end-stage renal disease (ESRD): patient on maintenance dialysis, including hemodialysis and peritoneal dialysis.

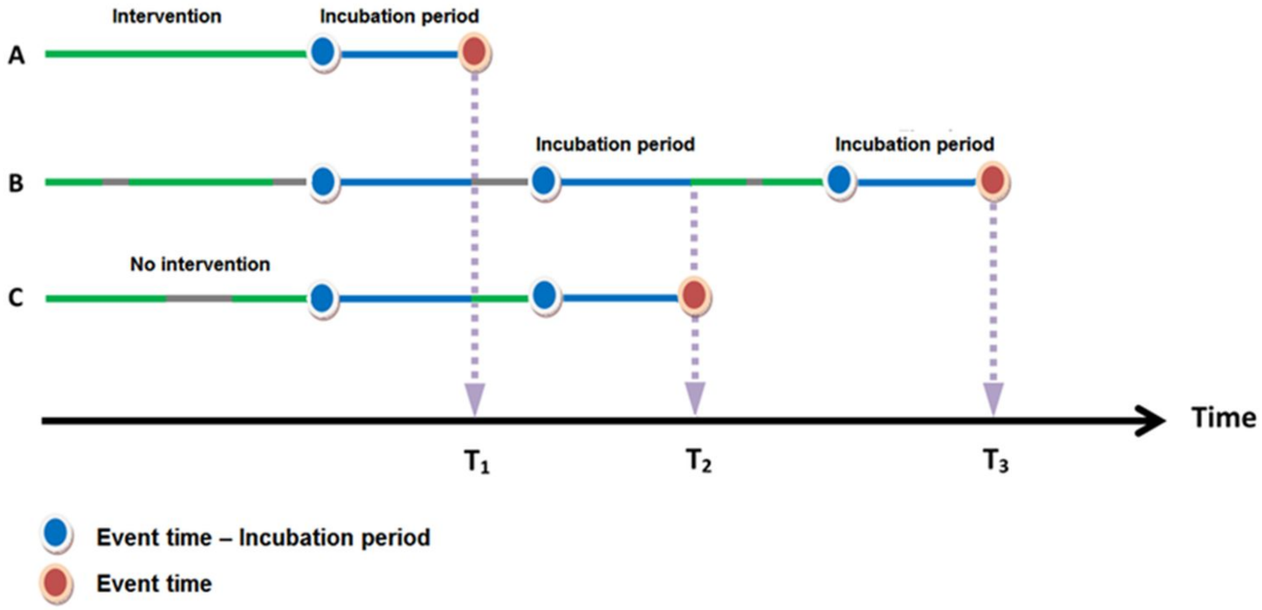
Telehealth Program

The fourth-generation telehealth program at our telehealth center is a synchronized and integrated remote management program for chronic conditions. The Internet-based platform was developed by the Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taiwan. The details of this program have been reported previously [15]. Briefly, this telehealth program provides the following services: (1) biometric data, including single-lead electrocardiography, blood pressure, heart rate, and pulse oximetry, which are transferred from patients to our telehealth center daily and on-demand; (2) nurse case managers communicate with patients daily and on-demand by telephone to ensure compliance to medication and medical instruction; (3) full-time nurse case managers and cardiologists are in charge of telehealth care 24 hours a day; and (4) long-term medications and management are discussed with the patients' primary care physicians if acute events occur. This telehealth program emphasizes compliance to medication and medical instruction, as well as prevention and early detection of clinical deterioration.

Data Collection

All demographic and clinical data were obtained from the electronic database of the hospital. The diagnosis of a chronic disease was based on the electronic database. The discharge diagnosis was used if there was disagreement between outpatient and discharge diagnoses. The endpoints of this study were all-cause and cardiovascular hospitalizations.

Figure 1. Incubation periods of the intervention’s effect at event times. The A, B, and C represent three fictitious participants. The green color indicates that the participant is using the telehealth service, the gray color indicates that the participant is not using telehealth service, and the blue color indicates that the participant is in the incubation period (defined as 28 days before a specific event). The T1, T2, and T3 are times when a participant develops an event (participants A, C, and B, respectively).



Statistical Analysis

Statistical analysis was performed using the R 3.3.1 software (R Foundation for Statistical Computing, Vienna, Austria). In statistical testing, a two-sided *P* value ≤.05 was considered statistically significant. The distributional properties of continuous variables were expressed as mean and standard deviation, categorical variables were presented as frequency and percentage, and the survival outcome curves were estimated using the Kaplan-Meier method. In the univariate analysis, the differences in the distributions of continuous variables, categorical variables, and survival outcomes were examined among three groups of patients (normal renal function, CKD, and ESRD) using the Kruskal-Wallis test, chi-square test, Fisher exact test, or log-rank test as appropriate for the data type. Next, multivariate analysis was conducted by fitting the Cox proportional hazards models with time-dependent covariates (called the “Cox model” for simplicity) to estimate the adjusted effects of risk factors, prognostic factors, and predictors on survival outcomes.

Contract Compliance Rate to the Telehealth Program

Contract compliance rate was defined as the percentage of days covered by the telehealth service duration within a certain time period before hospitalization. Specifically, we defined four time-varying variables of contract compliance for the time periods of 4, 8, 12, and 24 weeks, respectively, and then used them in fitting the Cox models. We assumed that the “incubation period” of the telehealth service was approximately 28 days; that is, the effect of the telehealth service might not appear until 28 days after signing or renewing the contract. The time-dependent contract compliance rates for 4, 8, 12, and 24 weeks were calculated using the following equation: contract compliance rates in *n* weeks = [(number of days covered by the telehealth program contract within *n* weeks before a specific event time – 28 days) / (7 days × *n*)] × 100%, where *n*=4, 8, 12, and 24. As shown in Figure 1, “a specific event time minus 28 days” is a specific event time moved 28 days backward to account for the assumed incubation period of the telehealth service.

We reorganized the original wide-form data into a long-form structure first, using the so-called “counting process style of input” for the survival outcome of interest, such as time to first hospitalization. Next, at each ordered event time of the survival outcome, we computed the values of the previously mentioned four time-dependent covariates for each of the patients at risk in the transformed long-form dataset. Finally, we fitted the Cox models to the long-form data with all relevant time-fixed and time-varying covariates using the coxph() function of the survival package in R.

To ensure the analysis quality, we used the model-fitting techniques for (1) variable selection, (2) goodness-of-fit assessment, and (3) regression diagnostics and remedies in our regression analyses. The stepwise variable selection procedure was applied to obtain the best candidate final regression model. The significance levels for entry and stay in the model were set to .15 to be conservative. The final regression model was identified manually by dropping the covariates with *P*>.05 one at a time until all regression coefficients were significantly different from zero.

Simple and multiple generalized additive models (GAMs) were fitted to detect nonlinear effects of continuous covariates and to identify appropriate cut-off point(s) for discretizing continuous covariates, if necessary, during the stepwise variable selection procedure. The vgam() function (with default values for the smoothing parameters) of the VGAM package was used to fit GAMs for continuous, binary, and count responses in R. Because GAMs were originally developed for smoothing the effects of continuous covariates in generalized linear models, we fitted GAMs of binary responses (eg, 1=hospitalization vs 0=no hospitalization) for our survival outcomes. In particular, we performed GAM analyses for the three groups of patients (normal, CKD, and ESRD) separately to detect heterogeneous nonlinear effects of continuous covariates in subgroups. Finally, the statistical tools of regression diagnostics for residual analysis, detection of influential cases, and check of multicollinearity were applied to discover any model or data problems. A variance-inflating factor value of 10 or greater in continuous covariates or 2.5 or greater in categorical covariates indicated the occurrence of multicollinearity problems among some of the covariates in the fitted regression model.

Results

Descriptive Statistics

In total, 715 patients were enrolled in this study and divided into three groups according to the status of their renal function: normal (n=490), with CKD (n=178), or with ESRD and on maintenance dialysis (n=47). The mean age was 66.6 (SD 15.0) years and 66.0% (472/715) of the participants were male. The patients' baseline characteristics, stratified by renal function status, are shown in Table 1.

Table 1. Baseline demography of patients stratified according to renal function status (N=715).

Baseline characteristics	Normal renal function (n=490)	Chronic kidney disease (n=178)	End-stage renal disease on dialysis (n=47)	*P*
Age (years), mean (SD)	63.7 (14.8)	74.8 (11.7)	69.7 (11.8)	<.001
Gender (male), n (%)	333 (68.0)	112 (62.9)	27 (57.4)	.15
Comorbidities, n (%)				
Hypertension	239 (48.7)	125 (70.2)	35 (74.5)	<.001
Diabetes	125 (25.5)	92 (51.7)	32 (68.1)	<.001
Cancer	57 (11.6)	20 (11.2)	4 (8.5)	.41
Atrial fibrillation	94 (19.2)	41 (23.0)	4 (8.5)	.07
Heart failure	130 (26.5)	80 (44.9)	30 (63.8)	<.001
Myocardial infarction	84 (17.1)	26 (14.6)	7 (14.9)	.79
Coronary artery disease	234 (47.8)	104 (58.4)	29 (61.7)	.07
Peripheral arterial disease	16 (3.3)	24 (13.9)	8 (17.0)	<.001
Cerebral vascular accident	50 (10.2)	32 (18.0)	6 (12.8)	.06
Hemodialysis	0	0	35 (74.5)	<.001
Peritoneal dialysis	0	0	12 (25.5)	<.001
Medications, n (%)				
Angiotensin-converting-enzyme inhibitor	43 (8.8)	12 (6.7)	3 (6.4)	.69
Angiotensin receptor blockers	197 (40.2)	84 (47.2)	17 (36.2)	.21
Beta-blocker	278 (56.7)	91 (51.1)	25 (53.2)	.45
Calcium channel blocker	149 (30.4)	87 (48.9)	22 (46.8)	<.001
Metformin	44 (9.0)	18 (10.1)	0	.12
Sulfonylurea	47 (9.6)	36 (20.2)	4 (8.5)	.001
Glitazones	3 (0.6)	6 (3.3)	0	.06
Dipeptidyl peptidase-4 inhibitor	41 (8.4)	33 (18.5)	12 (25.5)	<.001
Insulin	3 (0.6)	12 (6.7)	5 (10.6)	<.001
Spironolactone	55 (11.2)	34 (19.1)	0	<.001
Thiazide	42 (8.6)	24 (13.5)	2 (4.3)	.08
Statins	179 (36.5)	70 (39.3)	14 (29.8)	.69
Fenofibrate	5 (1.0)	17 (9.6)	2 (4.3)	<.001
Telehealth contract duration (days), mean (SD)	297 (410)	436 (506)	345 (407)	.001

Figure 2. Kaplan-Meier curve for time to the first cardiovascular (CV) hospitalization.

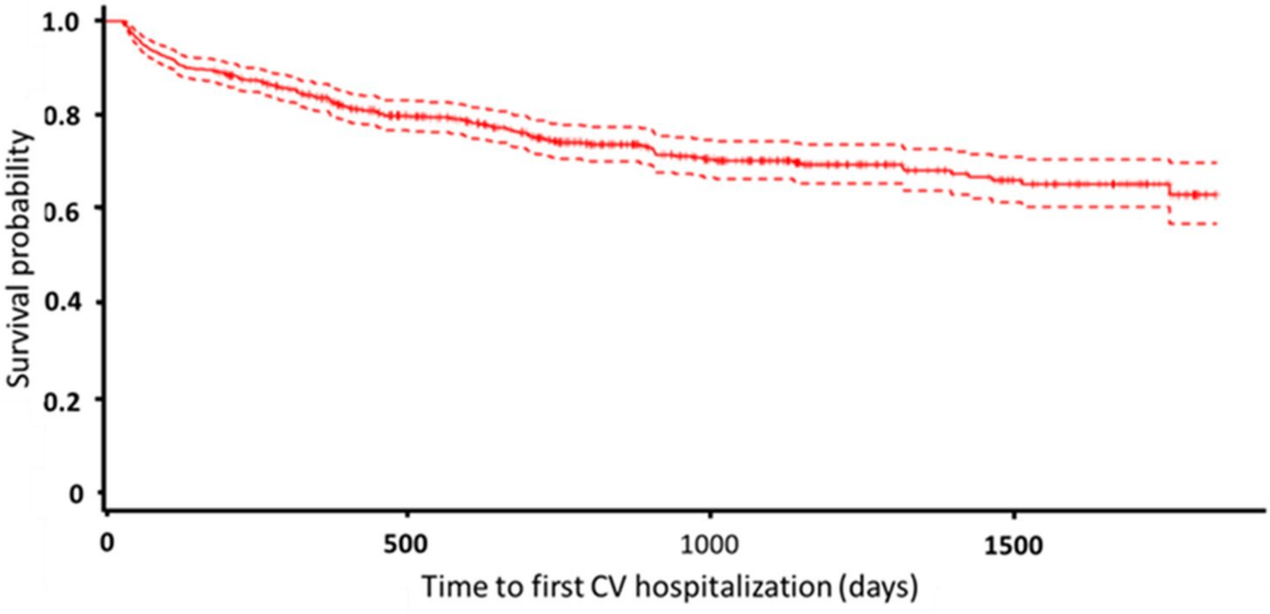
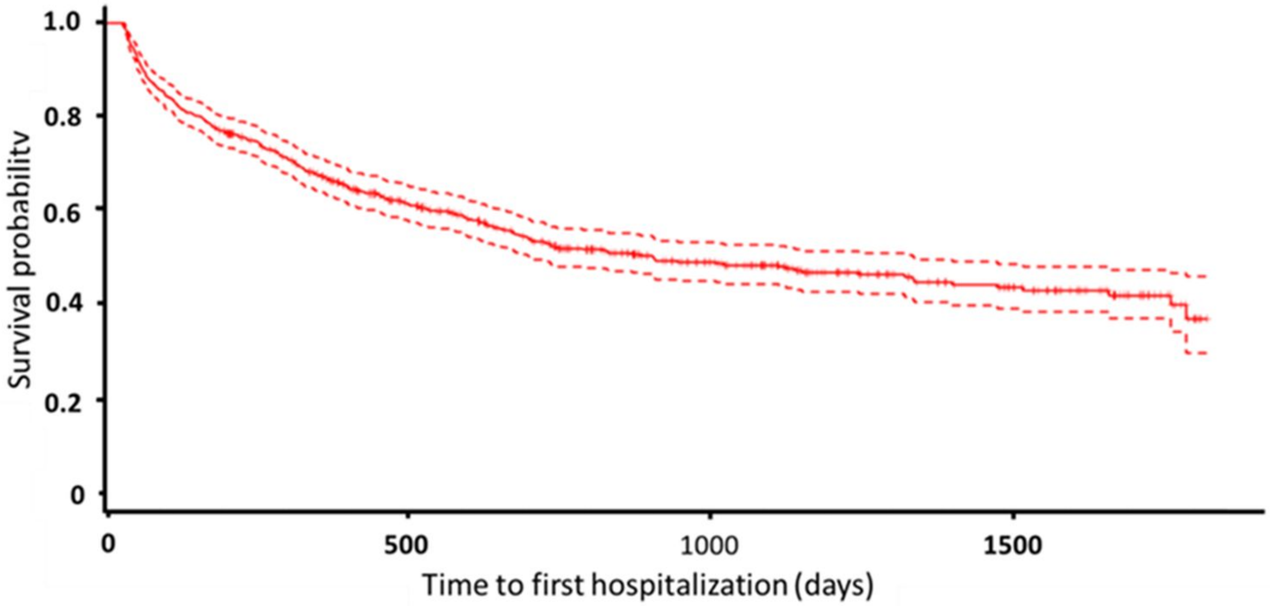


Figure 3. Kaplan-Meier curve for time to the first all-cause hospitalization.



The three groups were significantly different in terms of age and the proportion with hypertension, diabetes, heart failure, or peripheral artery disease, and the proportion receiving dialysis. The duration of telehealth use was different between the three study groups (longest in the group with CKD compared with the group with normal renal function and the group with ESRD on dialysis; Table 1). The median follow-up time was 694 days (IQR 338-1163). Because of the different follow-up times, events were divided according to follow-up time (days) in the subsequent analysis. During the follow-up period, there were 57 emergency room visits, 173 cardiovascular hospitalizations, and 350 all-cause hospitalizations. The Kaplan-Meier curves for time to first cardiovascular and all-cause hospitalization are shown in Figures 2 and 3, respectively.

Table 2. Univariate analysis of first cardiovascular hospital admission according to renal function status.

Variable	All patients		Normal renal function		Chronic kidney disease		End-stage renal disease on dialysis	
	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Male	1.14 (0.78-1.7)	.51	1.39 (0.82-2.42)	.22	0.93 (0.46-1.92)	.87	1.61 (0.40-6.73)	.54
Hypertension	0.93 (0.64-1.34)	.72	0.69 (0.42-1.12)	.13	0.74 (0.35-1.57)	.47	1.79 (0.34-10.5)	.48
Diabetes	1.74 (1.2-2.54)	.002	1.50 (0.88-2.53)	.11	1.15 (0.58-2.3)	.74	1.70 (0.36-8.57)	.51
Cancer	0.53 (0.26-1.01)	.05	0.61 (0.24-1.36)	.29	0.30 (0.05-1.10)	.08	0 (0.18-∞)	.49
Atrial fibrillation	1.06 (0.67-1.67)	.82	1.45 (0.80-2.56)	.18	0.61 (0.24-1.44)	.25	1.97 (0.09-124)	>.99
Heart failure	2.08 (1.43-3.03)	<.001	1.84 (1.09-3.08)	.02	1.61 (0.80-3.25)	.18	1.65 (0.38-7.56)	.52
Myocardial infarction	1.43 (0.89-2.27)	.12	1.58 (0.86-2.82)	.13	1.32 (0.49-3.42)	.65	1.32 (0.19-10.44)	>.99
Coronary artery disease	1.42 (0.99-2.06)	.05	1.59 (0.98-2.59)	.05	0.99 (0.49-1.99)	>.99	1.08 (0.26-4.54)	>.99
Peripheral artery disease	2.70 (1.36-5.34)	.003	2.36 (0.59-8.42)	.16	1.37 (0.48-3.75)	.63	7.27 (0.76-366.79)	.09
Cerebral vascular accident	1.15 (0.65-1.98)	.59	1.57 (0.72-3.24)	.25	0.72 (0.25-1.86)	.52	0.43 (0.03-3.46)	.41
Hemodialysis	3.13 (1.37-7.18)	.004	N/A	N/A	N/A	N/A	1.07 (0.23-5.09)	>.99
Peritoneal dialysis	2.82 (0.74-10.69)	.09	N/A	N/A	N/A	N/A	0.93 (0.20-4.44)	>.99
Angiotensin-converting-enzyme inhibitor	1.86 (1.00-3.42)	.03	2.14 (0.99-4.48)	.04	1.63 (0.37-6.75)	.51	1.97 (0.09-124)	>.99
Angiotensin receptor blockers	0.94 (0.65-1.36)	.79	1.07 (0.66-1.74)	.81	0.78 (0.39-1.57)	.51	0.75 (0.16-3.38)	.74
Beta-blocker	1.37 (0.94-1.98)	.94	1.47 (0.90-2.43)	.13	1.19 (0.60-2.39)	.63	1.97 (0.48-8.53)	.35
Calcium channel blocker	0.98 (0.67-1.43)	.67	0.69 (0.39-1.18)	.17	0.93 (0.47-1.87)	.87	1.61 (0.40-6.67)	.54
Metformin	2.1 (1.11-3.86)	.01	3.23 (1.53-6.72)	.002	1.06 (0.26-3.73)	>.99	N/A	N/A
Sulfonylurea	1.2 (0.65-2.02)	.58	1.11 (0.45-2.52)	.83	1.99 (0.40-2.37)	>.99	0.95 (0.06-14.39)	>.99
Glitazones	0.92 (0.09-5.2)	>.99	1.84 (0.03-35.75)	.51	0.47 (0.01-4.87)	.66	N/A	N/A
Insulin	0.98 (0.27-2.95)	>.99	0 (0-8.93)	>.99	0.61 (0.10-2.59)	.55	0.95 (0.06-14.39)	>.99
Spironolactone	2.13 (1.28-3.52)	.002	2.23 (1.12-4.34)	.02	2.23 (0.94-5.29)	.06	N/A	N/A
Thiazide	1.72 (0.95-3.06)	.94	1.87 (0.82-4.06)	.10	1.55 (0.56-4.16)	.35	0.95 (0.01-78.40)	>.99
Statins	1.52 (1.05-2.19)	.02	1.49 (0.91-.2.43)	.09	01.37 (0.35-1.48)	.40	0 (0.21-4.10)	>.99
Fenofibrate	1.16 (0.79-1.67)	.46	0 (0-4.01)	.59	0.73 (0.24-3.22)	.40	0 (0.18-∞)	.49

Univariate Analysis for First Cardiovascular Hospitalization

We performed a univariate analysis using each variable associated with the first cardiovascular hospitalization (Table 2). Among all study participants, a history of diabetes (OR 1.74, 95% CI 1.20-2.54, *P*=.002), heart failure (OR 2.08, 95% CI 1.43-3.03, *P*<.001), peripheral artery disease (OR 2.7, 95% CI 1.36-5.34, *P*=.003), or hemodialysis (OR 3.13, 95% CI 1.37-7.18, *P*=.004) were associated with a higher risk of first cardiovascular hospitalization. The use of angiotensin-converting-enzyme inhibitors (OR 1.86, 95% CI 1.00-3.42, *P*=.03), metformin (OR 2.10, 95% CI 1.11-3.86, *P*=.01), spironolactone (OR 2.13, 95% CI 1.28-3.52, *P*=.002), or statin (OR 1.52, 95% CI 1.05-2.19, *P*=.02) were also significantly associated with the first cardiovascular hospitalization. Among patients with normal renal function, only heart failure (OR 1.84, 95% CI 1.09-3.08, *P*=.02) was associated with a higher risk of first cardiovascular hospitalization. The associations between medication use and hospitalization in patients with normal renal function were similar to the overall population, except in the case of statins. However, there was no association between clinical factors or the use of medication and the first hospitalization in patients with CKD or ESRD.

Hazard Ratio for First Cardiovascular and All-Cause Hospitalization

We used Cox proportional hazards models with time-dependent covariates (ie, the Cox model) to estimate the effect of predictors on first cardiovascular hospitalization (Table 3). The results showed that previous emergency department admission, peripheral arterial disease, the use of spironolactone, the use of statins or metformin in patients with normal renal function, the presence of atrial fibrillation in patients with normal renal function, and the presence of ESRD were associated with a higher risk of first cardiovascular hospital admission. In addition, contract compliance rate of the telehealth program among patients with CKD was also associated with a higher risk of first cardiovascular hospital admission. We used a GAM plot to assess the nonlinear relationship between contract compliance rate within 24 weeks and risk of cardiovascular hospitalization (Figure 4). The results showed that there was a triphasic, or U-shaped, relationship: patients with very low or very high contract compliance rates within 24 weeks were associated with a higher risk of cardiovascular hospitalization.

The Cox model for estimating the effects of predictors on the first all-cause hospitalization yielded similar results (Table 4). The presence of cancer and the use of spironolactone were associated with a higher risk of first all-cause hospitalization, whereas the presence of normal renal function and the use of fenofibrate were associated with a lower risk of first all-cause hospitalization. We used a GAM plot to assess the nonlinear relationship between contract compliance rate within 24 weeks and risk of all-cause hospitalization. The results showed that there was a triphasic relationship: patients with very low and very high contract compliance rates within 24 weeks were associated with a higher risk of cardiovascular hospitalization in all three patient groups (normal renal function, CKD, and ESRD; Figures 5-7). Notably, patients with ESRD had a lower threshold for the right arm of the U-shaped relationship compared with the other two groups; in other words, their risk of all-cause hospitalization was higher if their contract compliance rate was greater than 44.4% compared with 89.1% and 91.8% for patients with normal renal function and CKD, respectively.

Table 3. Hazard ratio for first cardiovascular hospitalization.

Variable	Hazard ratio (95% CI)	*P* value
Pre-emergency department admission	7.6 (5.5-10.7)	<.001
Peripheral arterial disease	2.0 (1.3-2.7)	.006
Spironolactone	2.1 (1.7-8.5)	<.001
Normal renal function × statins	1.8 (1.4-3.2)	.002
Normal renal function × metformin	2.1 (1.2-3.5)	.007
Normal renal function × atrial fibrillation	1.6 (1.0-2.6)	.03
Chronic renal disease × 24-week contract compliance	2.5 (1.6-4.1)	<.001
End-stage renal disease	4.1 (2.5-6.9)	<.001

Figure 4. Generalized additive model (GAM) plot to assess the nonlinear relationship between contract compliance rate within 24 weeks and risk of cardiovascular hospitalization. Note: logit(P) was the logit transformation of probability(P)=ln(P/[1–P]).

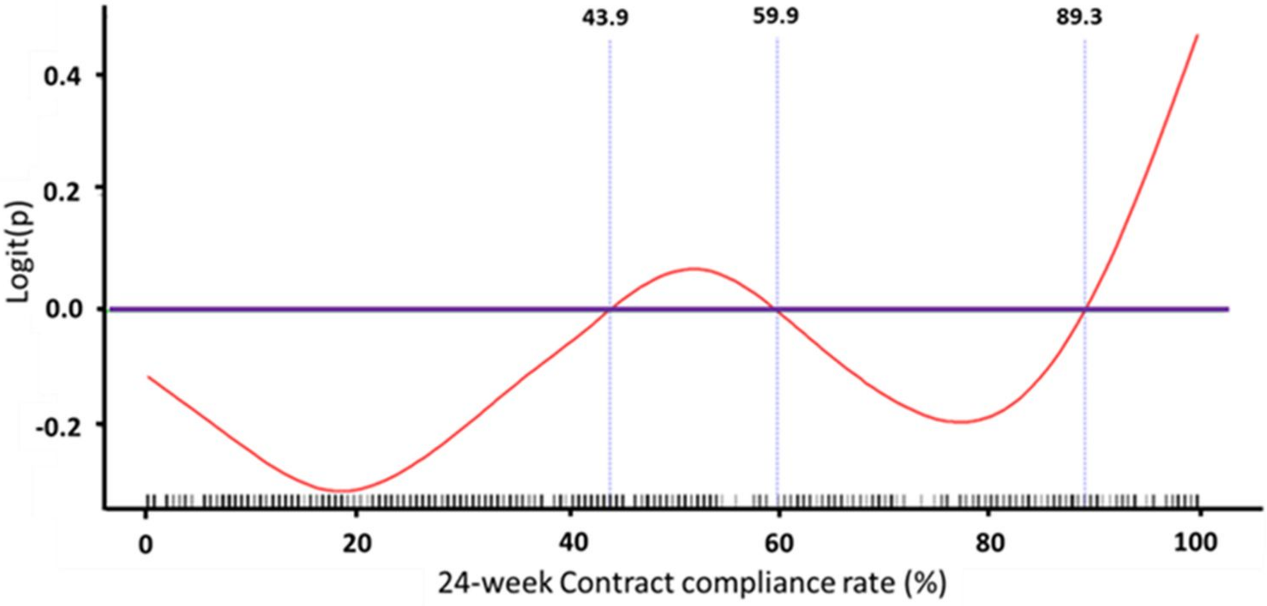


Table 4. Hazard ratio for first all-cause hospitalization.

Variable	Hazard ratio (95% CI)	*P*
Normal renal function	0.4 (0.3-0.5)	<.001
Cancer	1.5 (1.1-2.0)	.02
Fenofibrate	0.4 (0.2-0.9)	.03
Spironolactone	2.2 (1.6-2.9)	<.001
Normal renal function × 24-week contract compliance	2.2 (1.5-3.2)	<.001
Normal renal function × diabetes	1.6 (1.1-2.1)	.007
Normal renal function × cerebral vascular accident	1.7 (1.1-2.6)	.02
Normal renal function × peripheral arterial disease	2.1 (1.1-3.8)	.02
Chronic renal disease × 24-week contract compliance	2.0 (1.3-3.2)	.002

Figure 5. Generalized additive model (GAM) plot to assess the nonlinear relationship between contract compliance rate within 24 weeks and risk of cardiovascular hospitalization for participants with normal renal function. Note: logit(P) was the logit transformation of probability(P)=ln(P/[1–P]).

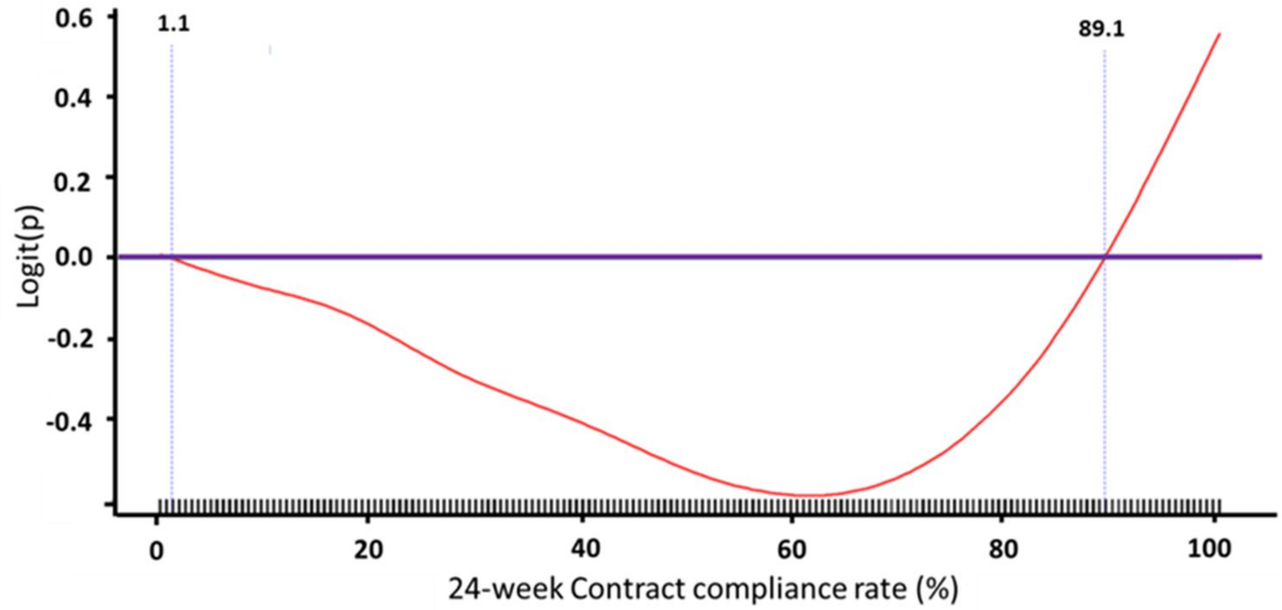


Figure 6. Generalized additive model (GAM) plot to assess the nonlinear relationship between contract compliance rate within 24 weeks and risk of cardiovascular hospitalization for participants with chronic kidney disease. Note: logit(P) was the logit transformation of probability(P)=ln(P/[1–P]).

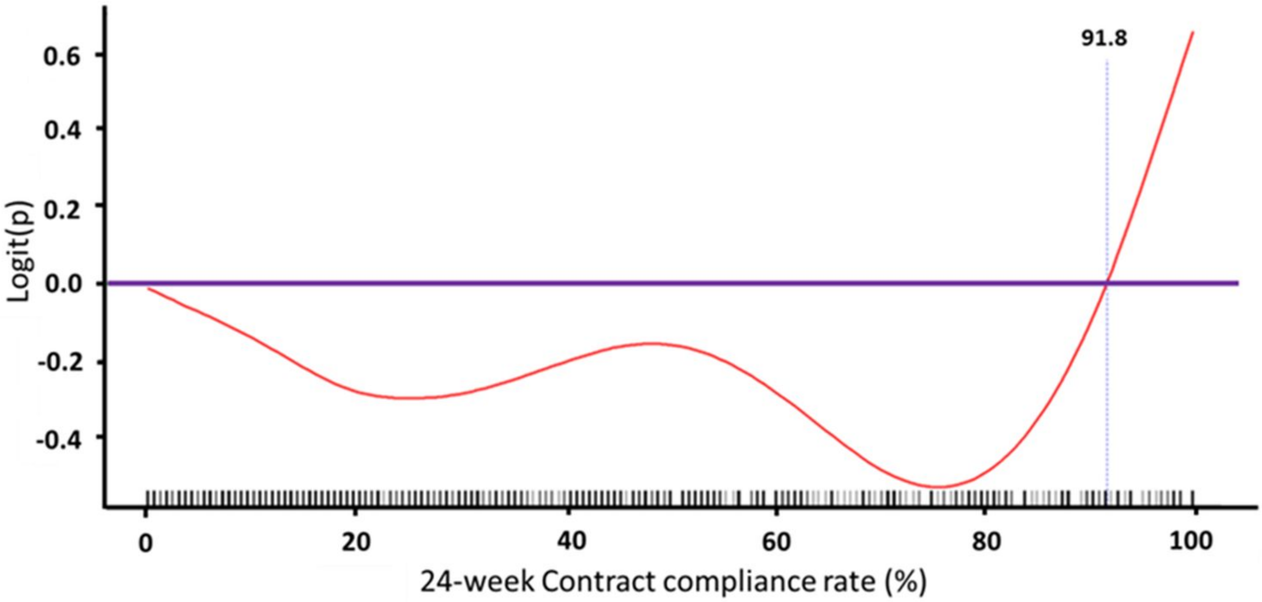
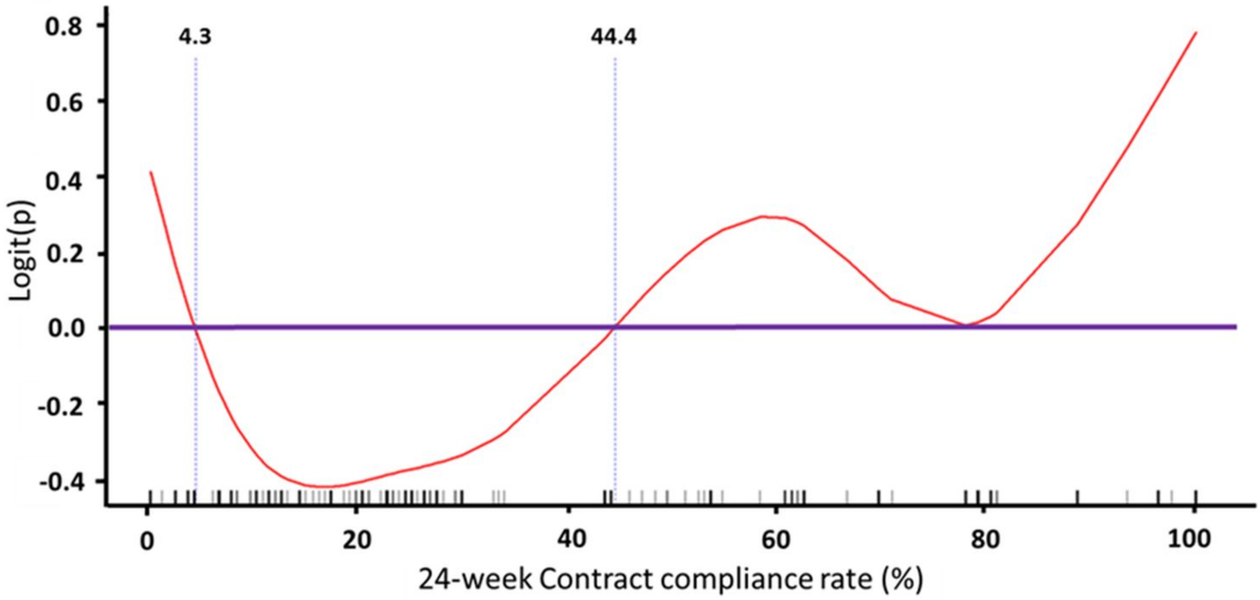


Figure 7. Generalized additive model (GAM) plot to assess the nonlinear relationship between contract compliance rate within 24 weeks and risk of cardiovascular hospitalization for participants with end-stage renal disease. Note: logit(P) was the logit transformation of probability(P)=ln(P/[1–P]).



Discussion

Major Findings

The major findings of this study were as follows: (1) higher contract compliance rate with the telehealth program was associated with a higher risk of first cardiovascular or all-cause hospitalization; (2) for patients receiving the telehealth program, the level of renal function was associated with hospitalization and patients with ESRD were associated with a higher risk of first cardiovascular or all-cause hospitalization after adjusting for comorbidities and medications; and (3) there was an interaction between the effect of contract compliance rate and renal function on the time to first hospitalization. Patients with ESRD had a higher risk of hospitalization at a lower contract compliance rate, compared with patients with CKD or normal renal function.

In our previous study, patients receiving the telehealth program were associated with lower all-cause mortality than those receiving usual care [5]. Based on that study and these findings, we suggest that one reason for this was timely hospitalization for disease management. Patients receiving the telehealth program and exhibiting high contract compliance rate may be more vulnerable because of complex underlying diseases and comorbidity, and if so would require comprehensive care, with which the telehealth program would assist. Our results highlight the effect of contract compliance rate in patients receiving the benefits of the telehealth program.

Contract Compliance Rate and Hospitalization

The analyses of contract compliance rate and risk of hospitalization revealed a triphasic relationship. Patients with a very low contract compliance rate were associated with a higher risk of hospitalization. This implies that duration of contract to the telehealth program of close to or less than 28 days has no effect on the prevention of acute deterioration in disease. Patients with a midrange contract compliance rate were associated with a lower risk of hospitalization. This highlights the usefulness of the telehealth program for most patients. Unexpectedly, patients with higher contract compliance were associated with a higher risk of hospitalization. This last finding seems counterintuitive because most compliance studies show that higher compliance to pharmacotherapy is associated with a lower risk of complications [16-19]. The issue of compliance, however, has not been formally explored in telehealth studies and should not be taken for granted [20]. The effect of compliance to a telehealth program is difficult, if not impossible, to evaluate using a randomized controlled trial. Therefore, careful analysis of observational data is needed to explore this issue. In randomized controlled trials, reported compliance is usually good, with up to 85% of participants still using the telehealth program until the end of the study [21,22]. However, the impact of contract compliance rate on the risk for hospitalization in the real world has not been reported. Our report shows that with a low to midrange contract compliance rate, a higher contract compliance rate is associated with a lower risk of hospitalization.

A similar U-shaped curve (the “three-phase terrain”) of hospital admission has been observed in lifetime analysis of heart failure [23,24]. In that report, 30% of all cardiovascular readmissions occurred within the first 2 months of hospital discharge, and 50% occurred within the 2 months before death, with much lower admission rates (15%-20%) observed in the intercurrent plateau phase [23]. To achieve a longer plateau phase, patient care should extend into the home to monitor and maintain stability, and into active intervention for ambulatory patients who have early signs of organ decompensation [25]. Given the predictive or preventive scenarios for heart failure admission, the relationship between high admission risk and compliance with telehealth demonstrated in this study merits further consideration. There are two possible explanations for this result. First, we consider the predictive scenario [25]. Because participation in the telehealth program is voluntary and self-funded, contract compliance rate is largely determined by the needs of the disease process and the sense that these needs have been met as a result of participation in the program. Therefore, patients with greater disease severity remain in the telehealth program for longer. However, greater disease severity is also associated with an increased risk of hospitalization. Second, to prevent further deterioration of the clinical situation, the telehealth program might partly increase the rate of hospitalization if a decline in organ function cannot be managed in outpatient or emergency departments. We have shown in our previous study that our program is associated with lower all-cause mortality rates. According to the preventive scenario, the program increases the rate of hospitalization so that crises can be managed earlier, rapid deterioration can be halted, and overall outcomes can be improved.

Renal Function and Hospitalization

Telehealth programs in CKD [26,27] have been less well studied compared with other chronic conditions such as heart failure, chronic obstructive pulmonary disease, or diabetes [2,28-31]. Knowledge about the efficacy and cost-effectiveness of telehealth among patients with CKD is still lacking. In a recent study, researchers showed that a telehealth program provided by an interprofessional team is a feasible approach for this patient group, but that it did not result in less hospitalizations, fewer emergency department visits, or lower all-cause mortality [26]. In our previous report, we showed that CKD was a predictor for all-cause mortality among these patients [5]. These results show that among patients receiving the telehealth program, renal function status was still a predictor for first hospitalization. Our program is specialized for risk factor control and adherence to recommended management guidelines among these patients. Determining whether telehealth improves the outcome of these patients requires further effectiveness comparison studies.

Interaction Between Renal Status and Contract Compliance Rate With Telehealth

We observed a significant interaction between the effects of renal status and contract compliance rate on hospitalization. Among patients with normal renal function and CKD, a contract compliance rate greater than 90% was associated with a higher rate of all-cause hospitalization, whereas among patients with ESRD, a contract compliance rate greater than 44% was associated with a higher rate of hospitalization. It is possible that contract compliance rate is a reflection of patient’s awareness about their disease severity. The significantly lower contract compliance rate associated with increased hospitalization rate in patients with ESRD might suggest that, among ESRD patients, their awareness may be lower than their real need for medical help. This could be due to factors such as low awareness of the severity and long-term complications of their disease, low awareness of the importance of compliance to recommended management guidelines, or accessibility of the health care system via alternative care providers (eg, the dialysis center). Whatever the case, noncompliance to recommended management guidelines in this patient group is associated with a higher risk of long-term complications [32,33]. From the perspective of telehealth care providers, it would be prudent to ascertain that patients with ESRD exhibit good compliance to standard management. Renal status, therefore, is a major determinant for risk evaluation of patients receiving telehealth care programs.

Limitations

There are several limitations in our study. Firstly, this is not a randomized controlled study comparing patients with and without access to the fourth-generation telehealth care program. Therefore, causation cannot be determined from our results. Secondly, there were fewer patients with CKD and ESRD (especially the latter) than with normal renal function. The effect of contract compliance rate was not statically significant for patients with ESRD, which may be due to limited patient numbers. Thirdly, our analysis did not account for recurrent events. It may be important to differentiate patients who are more likely to undergo several admissions from those with only one admission. Finally, our total patient number was limited. We have only listed the factors that were significant for predicting the first cardiovascular hospital admission in our final model; some predictors in the initial analyses were not statistically significant in the more comprehensive analyses (when the patients were divided into three groups). Thus, the interaction between clinical factors and each group may not be clear due to our limited patient numbers.

Conclusion

In summary, we found that contract compliance rate to the fourth-generation telehealth program is associated with cardiovascular outcomes, especially in patients with CKD. We demonstrated a triphasic relationship between contract compliance rate and risk of hospitalization among patients with CKD. Patients with low and high contract compliance rate to the telehealth program were associated with higher risk for hospitalization. A large, prospective, randomized controlled study is needed in these high-risk CKD patients to determine whether the telehealth program is beneficial.

Acknowledgments

We appreciated the financial support from the Ministry of Science and Technology of Taiwan (MOST 106-2314-B-002-164), and Center for Artificial Intelligence and Advanced Robotics, National Taiwan University (NTU-106R104062). We would also like to thank Uni-edit for editing and proofreading this manuscript.

Conflicts of Interest

None declared.

References

1. McLean S, Chandler D, Nurmatov U, Liu J, Pagliari C, Car J, et al. Telehealthcare for asthma. Cochrane Database Syst Rev 2010 Oct 06(10):CD007717. [doi: 10.1002/14651858.CD007717.pub2] [Medline: 20927763]
2. McLean S, Nurmatov U, Liu JL, Pagliari C, Car J, Sheikh A. Telehealthcare for chronic obstructive pulmonary disease. Cochrane Database Syst Rev 2011(7):CD007718. [doi: 10.1002/14651858.CD007718.pub2] [Medline: 21735417]
3. Clark RA, Inglis SC, McAlister FA, Cleland JG, Stewart S. Telemonitoring or structured telephone support programmes for patients with chronic heart failure: systematic review and meta-analysis. BMJ 2007 May 5;334(7600):942 [FREE Full text] [doi: 10.1136/bmj.39156.536968.55] [Medline: 17426062]
4. Inglis SC, Clark RA, McAlister FA, Ball J, Lewinter C, Cullington D, et al. Structured telephone support or telemonitoring programmes for patients with chronic heart failure. Cochrane Database Syst Rev 2010(8):CD007228. [doi: 10.1002/14651858.CD007228.pub2] [Medline: 20687083]
5. Hung C, Yu J, Lin Y, Chen Y, Huang C, Lee J, et al. Mortality benefit of a fourth-generation synchronous telehealth program for the management of chronic cardiovascular disease: a longitudinal study. J Med Internet Res 2016 May 13;18(5):e102 [FREE Full text] [doi: 10.2196/jmir.5718] [Medline: 27177497]
6. Go AS, Chertow GM, Fan D, McCulloch CE, Hsu C. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. N Engl J Med 2004 Dec 23;351(13):1296-1305. [doi: 10.1056/NEJMoa041031] [Medline: 15385656]
7. Wheeler DC, Becker GJ. Summary of KDIGO guideline. What do we really know about management of blood pressure in patients with chronic kidney disease? Kidney Int 2013 Mar;83(3):377-383 [FREE Full text] [doi: 10.1038/ki.2012.425] [Medline: 23325075]
8. Burnier M, Pruijm M, Wuerzner G, Santschi V. Drug adherence in chronic kidney diseases and dialysis. Nephrol Dial Transplant 2015 Jan;30(1):39-44. [doi: 10.1093/ndt/gfu015] [Medline: 24516224]
9. Remuzzi G, Benigni A, Finkelstein FO, Grunfeld J, Joly D, Katz I, et al. Kidney failure: aims for the next 10 years and barriers to success. Lancet 2013 Jul 27;382(9889):353-362. [doi: 10.1016/S0140-6736(13)60438-9] [Medline: 23727164]
10. Gallar P, Vigil A, Rodriguez I, Ortega O, Gutierrez M, Hurtado J, et al. Two-year experience with telemedicine in the follow-up of patients in home peritoneal dialysis. J Telemed Telecare 2007 Jun;13(6):288-292. [doi: 10.1258/135763307781644906] [Medline: 17785025]
11. Nakamoto H. Telemedicine system for patients on continuous ambulatory peritoneal dialysis. Perit Dial Int 2007 Jun;27 Suppl 2:S21-S26. [Medline: 17556306]
12. Kessler M, Erpelding M, Empereur F, Rose C. Transplantélic : une application de télémédecine en transplantation rénale. Eur Res Telemed 2013 Mar;2(1):5-10. [doi: 10.1016/j.eurtel.2013.01.002]
13. Wen CP, Cheng TY, Tsai MK, Chang YC, Chan HT, Tsai SP, et al. All-cause mortality attributable to chronic kidney disease: a prospective cohort study based on 462 293 adults in Taiwan. Lancet 2008 Jun 28;371(9631):2173-2182. [doi: 10.1016/S0140-6736(08)60952-6] [Medline: 18586172]
14. National Kidney Foundation. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. Am J Kidney Dis 2002 Feb;39(2 Suppl 1):S1-S266. [Medline: 11904577]

15. Ho Y, Yu J, Lin Y, Chen Y, Huang C, Hsu T, et al. Assessment of the cost-effectiveness and clinical outcomes of a fourth-generation synchronous telehealth program for the management of chronic cardiovascular disease. J Med Internet Res 2014;16(6):e145 [FREE Full text] [doi: 10.2196/jmir.3346] [Medline: 24915187]
16. Yao X, Abraham NS, Alexander GC, Crown W, Montori VM, Sangaralingham LR, et al. Effect of adherence to oral anticoagulants on risk of stroke and major bleeding among patients with atrial fibrillation. J Am Heart Assoc 2016 Feb 23;5(2):e003074 [FREE Full text] [doi: 10.1161/JAHA.115.003074] [Medline: 26908412]
17. Chen Y, Sloan FA, Yashkin AP. Adherence to diabetes guidelines for screening, physical activity and medication and onset of complications and death. J Diabetes Complications 2015;29(8):1228-1233 [FREE Full text] [doi: 10.1016/j.jdiacomp.2015.07.005] [Medline: 26316423]
18. Marzluf BA, Reichardt B, Neuhofer LM, Kogler B, Wolzt M. Influence of drug adherence and medical care on heart failure outcome in the primary care setting in Austria. Pharmacoepidemiol Drug Saf 2015 Jul;24(7):722-730. [doi: 10.1002/pds.3790] [Medline: 25980789]
19. Brown MT, Bussell JK. Medication adherence: WHO cares? Mayo Clin Proc 2011 Apr;86(4):304-314 [FREE Full text] [doi: 10.4065/mcp.2010.0575] [Medline: 21389250]
20. Hameed AS, Sauermann S, Schreier G. The impact of adherence on costs and effectiveness of telemedical patient management in heart failure: a systematic review. Appl Clin Inform 2014;5(3):612-620 [FREE Full text] [doi: 10.4338/ACI-2014-04-RA-0037] [Medline: 25298802]
21. Chaudhry SI, Mattera JA, Curtis JP, Spertus JA, Herrin J, Lin Z, et al. Telemonitoring in patients with heart failure. N Engl J Med 2010 Dec 9;363(24):2301-2309 [FREE Full text] [doi: 10.1056/NEJMoa1010029] [Medline: 21080835]
22. Takahashi PY, Pecina JL, Upatising B, Chaudhry R, Shah ND, Van HH, et al. A randomized controlled trial of telemonitoring in older adults with multiple health issues to prevent hospitalizations and emergency department visits. Arch Intern Med 2012 May 28;172(10):773-779 [FREE Full text] [doi: 10.1001/archinternmed.2012.256] [Medline: 22507696]
23. Chun S, Tu JV, Wijeysundera HC, Austin PC, Wang X, Levy D, et al. Lifetime analysis of hospitalizations and survival of patients newly admitted with heart failure. Circ Heart Fail 2012 Jul 01;5(4):414-421 [FREE Full text] [doi: 10.1161/CIRCHEARTFAILURE.111.964791] [Medline: 22556322]
24. Desai AS. The three-phase terrain of heart failure readmissions. Circ Heart Fail 2012 Jul 01;5(4):398-400 [FREE Full text] [doi: 10.1161/CIRCHEARTFAILURE.112.968735] [Medline: 22811548]
25. Desai AS, Stevenson LW. Rehospitalization for heart failure: predict or prevent? Circulation 2012 Jul 24;126(4):501-506 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.112.125435] [Medline: 22825412]
26. Ishani A, Christopher J, Palmer D, Otterness S, Clothier B, Nugent S, Center for Innovative Kidney Care. Telehealth by an interprofessional team in patients with CKD: a randomized controlled trial. Am J Kidney Dis 2016 Jul;68(1):41-49. [doi: 10.1053/j.ajkd.2016.01.018] [Medline: 26947216]
27. Thilly N, Chanliau J, Frimat L, Combe C, Merville P, Chauveau P, et al. Cost-effectiveness of home telemonitoring in chronic kidney disease patients at different stages by a pragmatic randomized controlled trial (eNephro): rationale and study design. BMC Nephrol 2017 Apr 05;18(1):126 [FREE Full text] [doi: 10.1186/s12882-017-0529-2] [Medline: 28381266]
28. Warren R, Carlisle K, Mihala G, Scuffham PA. Effects of telemonitoring on glycaemic control and healthcare costs in type 2 diabetes: a randomised controlled trial. J Telemed Telecare 2017 Jan 01:1357633X17723943. [doi: 10.1177/1357633X17723943] [Medline: 28814128]
29. Bashi N, Karunanithi M, Fatehi F, Ding H, Walters D. Remote monitoring of patients with heart failure: an overview of systematic reviews. J Med Internet Res 2017 Jan 20;19(1):e18 [FREE Full text] [doi: 10.2196/jmir.6571] [Medline: 28108430]
30. Lundell S, Holmner A, Rehn B, Nyberg A, Wadell K. Telehealthcare in COPD: a systematic review and meta-analysis on physical outcomes and dyspnea. Respir Med 2015 Jan;109(1):11-26 [FREE Full text] [doi: 10.1016/j.rmed.2014.10.008] [Medline: 25464906]
31. Faruque LI, Wiebe N, Ehteshami-Afshar A, Liu Y, Dianati-Maleki N, Hemmelgarn BR, Alberta Kidney Disease Network. Effect of telemedicine on glycated hemoglobin in diabetes: a systematic review and meta-analysis of randomized trials. CMAJ 2017 Mar 06;189(9):E341-E364 [FREE Full text] [doi: 10.1503/cmaj.150885] [Medline: 27799615]
32. Richard CJ. Self-care management in adults undergoing hemodialysis. Nephrol Nurs J 2006;33(4):387-394; quiz 395. [Medline: 17001996]
33. Mawar S, Gupta S, Mahajan S. Non-compliance to the continuous ambulatory peritoneal dialysis procedure increases the risk of peritonitis. Int Urol Nephrol 2012 Aug;44(4):1243-1249. [doi: 10.1007/s11255-011-0079-7] [Medline: 22102137]

Abbreviations

CKD: chronic kidney disease
CVD: cardiovascular disease
ESRD: end-stage renal disease
GAM: generalized additive model
GFR: glomerular filtration rate

Edited by G Eysenbach; submitted 07.09.17; peer-reviewed by H Barahimi, J Pecina, M Singh; comments to author 28.09.17; revised version received 10.11.17; accepted 18.12.17; published 24.01.18.

Please cite as:
Hung CS, Lee J, Chen YH, Huang CC, Wu VC, Wu HW, Chuang PY, Ho YL
Effect of Contract Compliance Rate to a Fourth-Generation Telehealth Program on the Risk of Hospitalization in Patients With Chronic Kidney Disease: Retrospective Cohort Study
J Med Internet Res 2018;20(1):e23
URL: http://www.jmir.org/2018/1/e23/
doi:10.2196/jmir.8914
PMID:29367185

Original Paper

Remote Collaborative Depression Care Program for Adolescents in Araucanía Region, Chile: Randomized Controlled Trial

Vania Martínez[1,2], MD, Dr med; Graciela Rojas[2,3], MD, Dr med; Pablo Martínez[2,4,5], PhD; Pedro Zitko[6,7], MD, MSc; Matías Irarrázaval[2,3,8], MD, MPH; Carolina Luttges[1,2], MSc; Ricardo Araya[9], MD, PhD

[1]Centro de Medicina Reproductiva y Desarrollo Integral del Adolescente, Facultad de Medicina, Universidad de Chile, Santiago, Chile

[2]Instituto Milenio para la Investigación en Depresión y Personalidad, Santiago, Chile

[3]Departamento de Psiquiatría y Salud Mental, Hospital Clínico Universidad de Chile, Santiago, Chile

[4]Centro de Innovación en Tecnologías de la Información para Aplicaciones Sociales, Universidad de Santiago de Chile, Santiago, Chile

[5]Escuela de Psicología, Facultad de Humanidades, Universidad de Santiago de Chile, Santiago, Chile

[6]Health Service & Population Research Department, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom

[7]Unidad de Estudios Asistenciales, Complejo Asistencial Barros Luco, Santiago, Chile

[8]Mental Health and Substance Use Unit, Pan American Health Organization/World Health Organization, Washington, DC, United States

[9]Centre for Global Mental Health and Primary Care Research, Health Service & Population Research Department, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, London, United Kingdom

Corresponding Author:
Graciela Rojas, MD, Dr med
Departamento de Psiquiatría y Salud Mental
Hospital Clínico Universidad de Chile
Av. La Paz 1003, Recoleta
Santiago,
Chile
Phone: 56 229788601
Email: graciela.rojas.castillo@gmail.com

Abstract

Background: Despite evidence on efficacious interventions, a great proportion of depressed adolescents do not receive evidence-based treatment and have no access to specialized mental health care. Remote collaborative depression care (RCDC) may help to reduce the gap between needs and specialized mental health services.

Objective: The objective of this study was to assess the feasibility, acceptability, and effectiveness of an RCDC intervention for adolescents with major depressive disorder (MDD) living in the Araucanía Region, Chile.

Methods: A cluster randomized, assessor-blind trial was carried out at 16 primary care centers in the Araucanía Region, Chile. Before randomization, all participating primary care teams were trained in clinical guidelines for the treatment of adolescent depression. Adolescents (N=143; 13-19 years) with MDD were recruited. The intervention group (RCDC, N=65) received a 3-month RCDC treatment that included continuous remote supervision by psychiatrists located in Santiago, Chile’s capital city, through shared electronic health records (SEHR) and phone patient monitoring. The control group (enhanced usual care or EUC; N=78) received EUC by clinicians who were encouraged to follow clinical guidelines. Recruitment and response rates and the use of the SEHR system were registered; patient adherence and satisfaction with the treatment and clinician satisfaction with RCDC were assessed at 12-week follow-up; and depressive symptoms and health-related quality of life (HRQoL) were evaluated at baseline and 12-weeks follow-up.

Results: More than 60.3% (143/237) of the original estimated sample size was recruited, and a response rate of 90.9% (130/143) was achieved at 12-week follow-up. A mean (SD) of 3.5 (4.0) messages per patient were written on the SEHR system by primary care teams. A third of the patients showed an optimal adherence to psychopharmacological treatment, and adolescents in the RCDC intervention group were more satisfied with psychological assistance than those in EUC group. Primary care clinicians were satisfied with the RCDC intervention, valuing its usefulness. There were no significant differences in depressive symptoms or HRQoL between groups. Satisfaction with psychological care, in both groups, was related to a significant change in depressive symptomatology at 12-weeks follow-up (beta=−4.3, 95% CI −7.2 to −1.3).

Conclusions: This is the first trial of its kind in Latin America that includes adolescents from vulnerable backgrounds, with an intervention that proved to be feasible and well accepted by both patients and primary care clinicians. Design and implementation issues may explain similar effectiveness across arms. The effectiveness of the intervention seems to be comparable with an already nationwide established treatment program that proved to be highly efficacious under controlled conditions.

Trial Registration: ClinicalTrials.gov: NCT01860443; https://clinicaltrials.gov/ct2/show/NCT01860443 (Archived by WebCite at http://www.webcitation.org/6wafMKlTY)

(J Med Internet Res 2018;20(1):e38) doi:10.2196/jmir.8021

KEYWORDS

primary health care; depression; adolescents; Internet; telemedicine; medically underserved area

Introduction

Adolescent Depression in Primary Care

At least 8% of adolescents attending primary care centers have probable major depressive disorder (MDD) [1]. Characterized by marked symptoms of low mood and anhedonia [2], adolescent MDD is associated with suicidal ideation [3,4], subsequent alcohol and illicit drug use [5], and poor living standards and worse mental health in adulthood [6,7]. Despite the burden of adolescent depression and the availability of efficacious interventions [8-10], a great proportion of adolescents with MDD do not receive evidence-based treatment or get no treatment at all [11,12].

Primary care has a major role in the delivery of depression treatment for adolescents [13]. The collaborative care model, defined as the team-driven, population-focused, measurement-guided, and evidence-based provision of care [14], has been regarded as one of the best approaches in integrated behavioral health [15]. Its effectiveness in treating adolescents with MDD has also been demonstrated, achieving significant improvements in depressive symptoms and global functioning [16,17]. However, most evidence comes from trials conducted in the United States [16-20], with multifaceted interventions to provide mental health care to mostly non-Hispanic white and urban adolescents [17-20].

The Chilean Case: the Use of Technologies and Collaborative Care for Depression

Chile, a developing country with a growing income inequality [21], has made substantial efforts during the past decades to improve mental health services, providing one of the first examples of an evidence-based depression intervention being scaled up in resource-constrained settings [22,23]. Yet, primary care providers' barriers to treating depression [24] coupled with the uneven geographical distribution of mental health resources, which especially affects low-income people living in hard-to-reach areas [25], make the Chilean case worth studying.

Being a country with an above-average digital connectivity [26], the use of information and communication technologies (ICTs) in Chile may help overcome the inequity of access to mental health resources. According to international evidence, ICTs have made it possible to effectively assist underserved populations through remote collaborative care programs [27,28], obtaining results comparable with face-to-face interventions in the management of depression that have mostly benefited white rural adults [29,30]. No adolescent-specific electronic mental health collaborative care for depression has been tested before.

This study tested the feasibility, acceptability, and effectiveness of a remote collaborative care program delivered by primary and specialized mental health teams to enhance the management of adolescent MDD in the Araucanía Region, an underserved area of Chile.

Methods

Study Design

This was an assessor-blind, 2-group, cluster randomized (1:1 ratio) clinical trial, carried out in the Araucanía Region, Chile. Randomization was conducted using computer-generated random numbers.

Setting, Participants, and Eligibility Criteria

Araucanía Region is the poorest administrative division of Chile (20.7% of households living below the poverty line) and has the highest proportion of ethnic minorities and rural population [31,32]. Compared with the national average, children and adolescents of the region live in precarious material conditions and suffer from the highest suicide rates, with 6.7 deaths due to suicide per 100,000 inhabitants [32]. Temuco, the capital city of Araucanía Region, is located 380 miles south of Santiago, Chile's capital city.

This clinical trial was conducted in 16 primary care centers located in the Araucanía Region. These centers were required to have an Internet connection, receive no psychiatric consulting services for children or adolescents, and ensure the recruitment of 15 depressed adolescents over a 13-month period. The health centers entered the study after being authorized by their boards of directors and receiving the informed consent of at least one family physician and one psychologist from each health team.

Adolescents aged between 13 and 19 years with suspected depression were invited (face-to-face) to participate in the study by their primary care teams; the informed consent of those over 18 years and the informed assent of minors along with their parents' or primary caregivers' informed consent were obtained. A child and adolescent psychiatrist in Santiago verified over the telephone that the adolescents could be diagnosed with depression and not psychotic depression, bipolar disorder, comorbidity with substances and or or alcohol dependence, or suicide risk requiring immediate specialized treatment. To do this, the semistructured clinical interview MINI-KID

(Mini-International Neuropsychiatric Interview For Children and Adolescents) was used, which enables researchers to make a diagnostic assessment in children and adolescents according to the criteria of International Statistical Classification of Diseases and Related Health Problems, 10th Revision, and Diagnostic and Statistical Manual of Mental Disorders, 4th Edition [33].

Interventions

Computer-Based Remote Depression Training for Health Teams

Before the random assignment, all the health teams received a Web-based training to improve their early detection of depression in adolescents and increase their access to timely and effective treatment. The training program was designed by 2 psychiatrists, 1 psychologist, and 1 social worker from Universidad de Chile's Faculty of Medicine. It was aimed at presenting a comprehensive approach to depression and suicide in adolescence and a critical assessment of the key recommendations issued by the Chilean Ministry of Health to overcome problems usually encountered in clinical practice when treating depression in adolescents.

This program was delivered over 30 hours, during which teaching activities were conducted using audiovisual material to present clinical cases. In addition, the participating health teams were encouraged to participate actively in discussion forums. The participants' learning outcomes were assessed with online written worksheets in which they were requested to apply their clinical judgment to tackle the clinical cases presented earlier and a multiple-choice test that considered the material covered in the program.

Remote Collaborative Depression Care for Adolescents

A remote collaborative depression care (RCDC) program for adolescents introduced shared electronic health records (SEHR) to improve communication between the clinicians in the Araucanía Region and the team of child and adolescent mental health specialists of Universidad de Chile's Faculty of Medicine, operating via an Internet platform in a secure virtual environment provided by the faculty. The SEHR functioned as a discussion forum allowing for personalized, confidential, and real-time interaction between the primary care teams and the specialists to assist the former during the diagnostic process and the treatment of the acute phase of the disorder.

In addition, after each patient's data was entered into his or her SEHR, a structured phone monitoring system hosted by Universidad de Chile's Faculty of Medicine was implemented by a psychologist trained in Santiago. Phone calls were done at 1, 2, 3, 6, and 9 weeks post baseline assessment, each lasting for about 5 to 10 min per adolescent and his/her primary caregiver. Contents of this component included monitoring of symptoms and adherence to pharmacological treatment, assessment of side effects of medications, evaluation of the quality of the therapeutic relationship, and agreements on the goals and tasks of the therapy. If patients could not be contacted, several attempts were made on different days of the week and times of the day, with no restrictions imposed on the number of attempts. This made it possible to triangulate the information obtained from the primary care teams, thus providing the specialists with other elements to be included in the online assistance sessions.

Enhanced Usual Care

The clinicians in the Araucanía Region centers that were randomly assigned to the control group were encouraged to follow the recommendations of the Ministry of Health of Chile for the management of adolescent depression and also received the main conclusions of the baseline diagnostic assessment.

Assessments

The feasibility of the project was assessed in terms of achieving the recruitment targets, the attrition rates at 12 weeks, and the use of SEHR by the primary care teams.

The acceptability of the intervention was assessed by comparing rates of pharmacological adherence and user satisfaction with the components of the treatment displayed by each group at 12-week follow-up. In addition, data reflecting the satisfaction of the primary care teams with the RCDC were analyzed.

Outcomes regarding the acceptability of the intervention were assessed as follows: Pharmacological adherence was assessed with a brief structured interview, asking patients whether their physician prescribed medications and whether those medications were taken correctly during the last 4 days and, if not, with the reasons for nonadherence. User satisfaction with treatment was evaluated through a self-reported questionnaire with a 7-point Likert scale, which included 5 items to rate treatment, facilities, medical care, psychological care, and nonprofessional staff treatment. For clinicians in the active group, the use of SEHR was noted and an RCDC satisfaction questionnaire was administered.

The primary outcome with respect to the effectiveness of the RCDC was assessed with mean Beck Depression Inventory (BDI) [34] scores at 12 weeks after treatment started. The BDI is a self-reported questionnaire for assessing depressive symptoms in people aged 13 years and older [34], which has been previously used in Chile [35]. The secondary outcome was change in health-related quality of life (HRQoL) at 12 weeks measured through the KIDSCREEN-27 questionnaire [36]. The KIDSCREEN-27 questionnaire measures the frequency or intensity of behaviors, feelings, or attitudes of the adolescent about different aspects of his/her everyday life during the previous week [36]. This questionnaire has been validated in Chile [37].

Patient baseline data and outcomes at 12-week follow-up were evaluated via telephone by a trained consultant at Universidad de Chile's Faculty of Medicine, who was blinded to treatment allocation.

Statistical Analysis

Sample Size

On the basis of previous studies [35,38], a sample size of 237 depressed adolescents was needed to detect a difference between groups in terms of an expected reduction in mean Beck scores of 3.2 points across arms at 12 weeks, representing approximately one-third of SD difference. This sample size has a power of 82%, with a one-sided alpha of .05.

Data Analysis

Descriptive analyses for relevant variables were conducted for the total sample and for each group separately to assess the balance across arms after randomization. Fisher exact test and independent samples *t* test were used to compare differences in patient adherence and treatment satisfaction between groups. Primary and secondary outcomes are described in terms of mean and SD pre- and postintervention. The magnitude of the effect of the intervention was evaluated using regression models, which included a random effect to account for intracluster correlations within primary care centers. Univariate and multivariate models were employed. Multivariate models included the baseline outcome value and variables that were not well balanced after randomization. Furthermore, using the same regression models, we explored the relevance of other health care features on the magnitude of change on the primary outcome. Regression coefficients are presented with their respective 95% CIs. Because missing primary outcome values were less than 7%, a sensitivity analysis was not conducted. All analyses were conducted using R 3.1.1 (R Foundation for Statistical Computing, Vienna, Austria) and its package lme4 for multilevel analysis. This trial has been reported in accordance to the CONSORT-EHEALTH guidelines [39].

Ethics

The study was approved by the Committee of Human Research Ethics of Universidad de Chile's Faculty of Medicine (Nº 080-2012), the Committee of Scientific Ethics of the North Araucanía Health Service, and the Committee of Scientific Ethics of the South Araucanía Health Service. The study allowed for participants' voluntary withdrawal with no adverse consequences. The local teams treating the participants, both in the active and in the control groups, were informed of cases in which high suicide risk or any other high-risk situations were detected for them to carry out the necessary actions.

Results

Sample Characteristics

The sample comprised 143 adolescents with MDD. Most of the participants were women (81.1%), aged 15.4 years (SD 1.6), living with both parents (47.6%), and of middle-low socioeconomic status (45.5%). Of the participants, 46.9% reported having had at least one mental health problem in their life, whereas 37.8% reported the same regarding one of their close relatives. Table 1 shows the baseline sociodemographic characteristics and clinical information of the sample, stratified by treatment group (enhanced usual care or EUC or RCDC). Significant differences exist between the groups in terms of socioeconomic status (*P*=.03).

Feasibility

Recruitment and Follow-Up

In total, 16 primary care centers participated in the study; one of these withdrew and was replaced. The participating health care centers were able to identify 178 adolescents with suspected depression and managed to recruit 143 eligible individuals, which represents 60.3% of the initially projected sample size. Of the participants, 45.5% received the RCDC intervention, whereas 54.5% received EUC (Figure 1).

At 12-week follow-up, 5 cases (6.4%) in EUC were lost to follow-up, and there was no lost to follow-up in RCDC (Figure 1).

Use of Shared Electronic Health Records

The team of child and adolescent mental health specialists wrote on the Internet platform a mean (SD) of 9.9 (4.2) times per patient, ranging from 3 to 28 messages, throughout the study period (12 weeks). The experts helped by providing continuous training for the health staff, general recommendations about actions and strategies to be implemented, and specific case management guidelines. They covered the following areas: assessment of depressive symptoms, medical treatment of depression (pharmacological and nonpharmacological), and psychosocial interventions. Their specific indications covered aspects such as confidentiality and its limits, sexual and reproductive health counseling, and working with family and school support networks.

Regarding the primary care teams, it was observed that those in the RCDC group wrote a mean (SD) of 3.5 (4.0) messages per patient. In 13 of the cases assigned to the RCDC group, 20% of the adolescents receiving RCDC, the clinicians from primary care teams wrote no messages, whereas in 16.7% of cases they did so only once.

Acceptability

Patient Adherence and Treatment Satisfaction

In general, nearly one-third of the sample displayed a regular intake of the psychoactive medications prescribed (n=44). No significant differences were observed across groups at 12-week follow-up in terms of adherence to pharmacological treatment (Fisher exact test *P* value=.98). Satisfaction outcomes reached 6 points on a scale from 1 to 7, with no differences across arms. Satisfaction with psychological care was the only aspect of treatment satisfaction that displayed a significant difference between the groups (Wilcoxon rank-sum test *P* value=.04), with the RCDC intervention scoring higher (Table 2).

In addition, patients in the active group received a mean (SD) of 3.8 (1.0) monitoring phone calls lasting 5 to 10 min each. On a scale from 1 (worst) to 7 (best), patients assigned a mean (SD) score of 6.2 (1.2) to the usefulness of these calls and 6.2 (1.2) to how comfortable they felt with the phone monitoring.

Table 1. Baseline characteristics of the sample.

Variables	EUC[a] (N=78)	RCDC[b] (N=65)	*P* value for between-groups differences
Sex (female), mean (SD)	64 (82.1)	52 (80.0)	.83[c]
Age in years, mean (SD)	15.6 (1.7)	15.2 (1.5)	.20[d]
Mapuche ethnicity[e], n (%)	11 (14)	13 (20)	.38[c]
Living in a rural area, n (%)	11 (14)	14 (22)	.27[c]
Living with, n (%)			.88[c]
Both parents	34 (44)	34 (52)	
Mother only	22 (28)	16 (25)	
Mother and her partner	8 (10)	5 (8)	
Father only	4 (5)	5 (8)	
Father and his partner	1 (1)	0 (0)	
Other relatives	8 (10)	5 (8)	
Other nonrelatives	1 (1)	0 (0)	
Socioeconomic status, n (%)			.03[c]
Low	11 (14)	16 (25)	
Mid-low	33 (42)	32 (49)	
Middle	22 (28)	16 (25)	
Mid-high	8 (10)	1 (2)	
High	4 (5)	0 (0)	
Years of schooling, mean (SD)	9.3 (2)	8.9 (2)	.10[d]
Personal history of mental illness, n (%)	33 (42)	34 (52)	.24[c]
Family history of mental illness, n (%)	34 (44)	20 (31)	.12[c]
Beck Depression Inventory score, mean (SD)	27.1 (9)	27.8 (10)	.65[f]
Perceived health[g], mean (SD)	3.6 (1)	3.6 (1)	.97[d]
Physical well-being[h], mean (SD)	30.9 (9)	31.4 (9)	.78[f]
Psychological well-being[h], mean (SD)	29.6 (10)	30.2 (10)	.98[d]
Autonomy and parents[h], mean (SD)	37.6 (8)	39.1 (9)	.67[f]
Peers and social support[h], mean (SD)	40.1 (13)	40.3 (14)	.92[f]
School environment[h], mean (SD)	37.8 (8)	38.2 (8)	.82[f]
Health-related quality of life index[i], mean (SD)	33.8 (7)	33.5 (7)	.84[f]

[a] EUC: enhanced usual care.

[b] RCDC: remote collaborative depression care.

[c] Fisher exact test *P* value.

[d] Wilcoxon rank-sum test *P* value.

[e] Mapuche are the indigenous people of the region.

[f] Student *t* test for unequal variances.

[g] KIDSCREEN-27 item score.

[h] KIDSCREEN-27 dimension score.

[i] Uses 10 items derived from the 27-item version of KIDSCREEN.

Figure 1. Flow diagram. EUC: enhanced usual care; PCCs: primary care centers; RCDC: remote collaborative depression care.

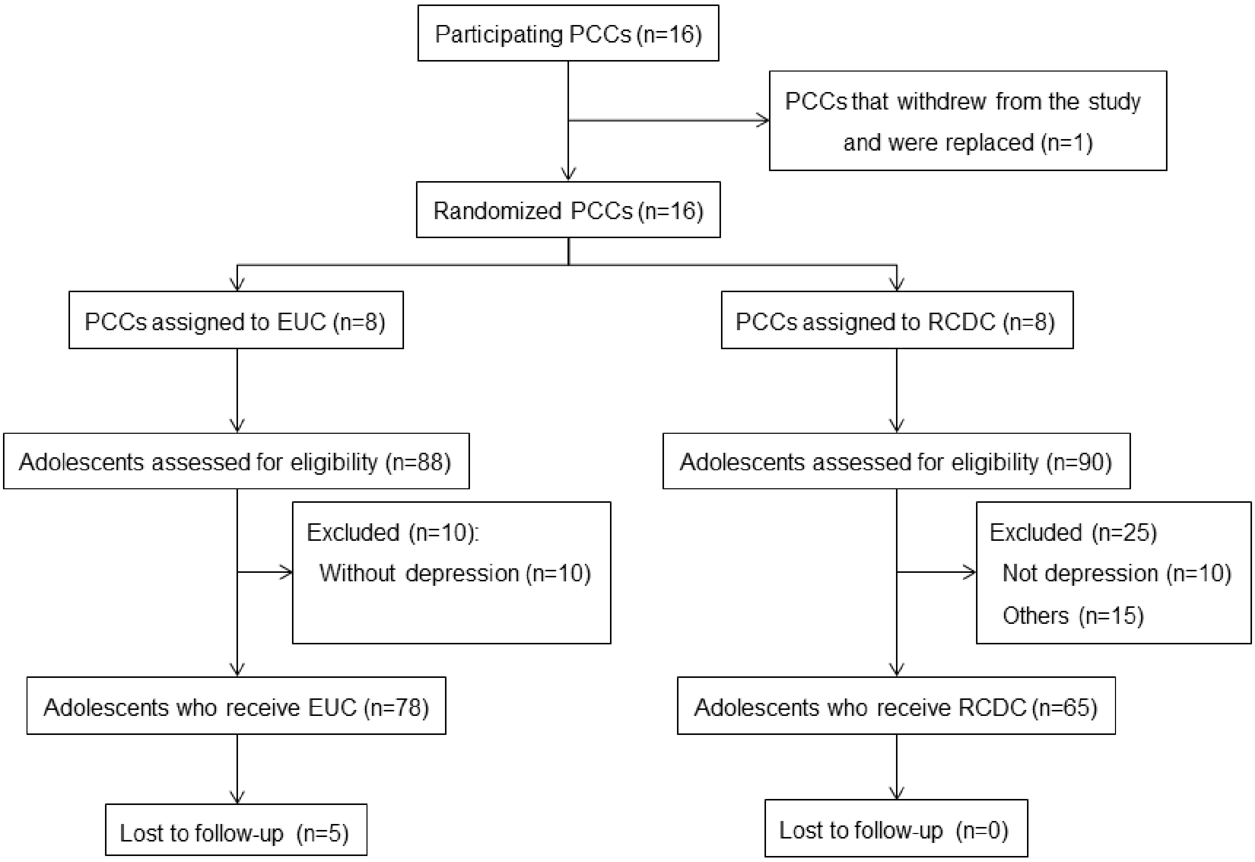


Table 2. Patient satisfaction with treatment at 12-week follow-up in the remote collaborative depression care (RCDC) and the usual care group.

Outcomes	EUC[a]		RCDC		
	N[b]	Mean[c] (SD)	N[b]	Mean[c] (SD)	Wilcoxon rank-sum test *P* value
Satisfaction with treatment	70	6.0 (1.1)	63	6.2 (1.2)	.19
Satisfaction with facilities	70	6.2 (1.0)	63	6.3 (1.0)	.45
Satisfaction with medical care	70	6.4 (1.2)	63	6.7 (0.8)	.08
Satisfaction with psychological care	69	6.4 (0.9)	61	6.7 (0.7)	.04
Satisfaction with nonprofessional staff treatment	69	6.2 (1.1)	63	6.4 (0.9)	.65

[a] EUC: enhanced usual care.

[b] Cases providing complete data and analyzed.

[c] Possible scores ranging from 1 (worst) to 7 (best).

Clinician Satisfaction With Remote Collaborative Depression Care

The primary care teams in the active group assessed the following aspects of the RCDC intervention: (1) usefulness for clinical work—mean (SD) score 6.3 (1.0), ranging from 3 to 7 on a scale from 1 (worst) to 7 (best), mode 7; (2) usefulness for patients—mean score 6.5 (0.7), range 4 to 7, mode 7; and (3) comfort level of the Internet platform—mean score 6.0 (1), range 4 to 7, mode 7.

In addition, the main positive aspects, difficulties, and suggestions regarding the RCDC voiced by the primary care clinicians in the active group are summarized in Textbox 1.

Effectiveness

Depressive Symptoms and Health-Related Quality of Life

No significant differences were observed across arms at 12-week follow-up in terms of depressive symptomatology or HRQoL before and after adjusting for baseline values and when introducing socioeconomic variables (Table 3).

Textbox 1. Main positive aspects, difficulties, and suggestions regarding the remote collaborative depression care (RCDC) intervention stated by clinicians in the active group.

Positive aspects

- The preintervention training was useful.
- The Internet platform is an innovative resource.
- It is useful to receive timely specialized support.
- The phone assessment and monitoring of the patients contributes to their positive perception of the treatment and fosters their trust in the local health team.

Difficulties

- Clinicians’ times for accessing the Internet platform were not scheduled
- Lack of interest in mental health displayed by some physicians
- High turnover of professionals in local health care centers
- Depressed adolescents tend to resist interventions
- Insufficient communication with patients and their parents in phone monitoring
- Very complex cases with limited follow-up time
- Medical work receives more support than psychotherapeutic work

Suggestions

- Ensuring the continuity of interventions of this type
- Better coordination between local and specialized teams
- Facilitating the internal coordination of the local team
- More user-friendly Internet platform
- Conducting face-to-face or videoconference assistance sessions
- Assigning and scheduling local equipment use for local teams to fulfill tasks required by RCDC
- Fostering the participation of the whole health care team in distance assistance sessions
- Incorporating email alerts to be sent to clinicians

Table 3. Magnitude of the effect of the intervention on depressive symptoms and health-related quality of life.

Outcomes	Analyzed[a], n (%)	Nonadjusted difference between means (95% CI)[b]	Adjusted difference between means[c] (95% CI)
Beck Depression Inventory score	138 (0.97)	1.1 (−2.9 to 5.2)	1.5 (−2.4 to 5.6)
Perceived health[d]	134 (0.94)	−0.1 (−0.4 to 0.2)	−0.2 (−0.4 to 0.1)
Physical well-being[e]	134 (0.94)	1.7 (−2.1 to 5.4)	2.3 (−0.8 to 5.7)
Psychological well-being[e]	134 (0.94)	0.8 (−3.9 to 5.8)	1.0 (−3.4 to 5.8)
Autonomy and parents[e]	134 (0.94)	−0.7 (−4.2 to 2.5)	−0.1 (−3.1 to 2.9)
Peers and social support[e]	134 (0.94)	−1.3 (−5.6 to 3.5)	−0.2 (−4.3 to 3.8)
School environment[e]	109 (0.76)	−3.5 (−7.4 to 0.7)	−3.1 (−7.1 to 1.0)
Health-related quality of life index[f]	110 (0.77)	1.0 (−2.8 to 4.8)	1.2 (−3.2 to 5.2)

[a] Number of patients who provided outcome data at 12-week follow-up and percentage of full outcome data.

[b] Effect magnitude estimated with Markov Chain Monte Carlo procedures.

[c] Adjusted according to the baseline measurement and socioeconomic status differences observed.

[d] KIDSCREEN-27 item score.

[e] KIDSCREEN-27 dimension score.

[f] Uses 10 items derived from the 27-item version of KIDSCREEN.

Table 4. Adjusted regression model for change in depressive symptomatology at 12-week follow-up (N=130).

Variables	Depressive symptomatology[a] at 12 weeks (95% CI)
Constant	32.0 (2.8-61.1)
Treatment group[b]	−8.0 (−40.0 to 24.0)
Initial depressive symptomatology[c]	−0.5 (−0.7 to −0.3)
Age	−0.1 (−1.4 to 1.2)
Sex	−0.1 (−5.3 to 5.1)
Ethnicity	−1.0 (−6.6 to 4.6)
Rurality	1.1 (−4.4 to 6.6)
Satisfaction with the psychological care received	−4.3 (−7.2 to −1.3)
Satisfaction with the psychological care received by treatment group	1.7 (−3.1 to 6.5)

[a] Beck Depression Inventory score at 12-week follow-up.

[b] Enhanced usual care or remote collaborative depression care.

[c] Baseline Beck Depression Inventory score.

Model for Change in Depressive Symptoms at 12-Week Follow-Up

Changes in depressive symptomatology at 12-week follow-up were explored with a regression model that considered the group to which the participants belonged, their baseline depressive symptomatology (according to BDI score), sex, age, ethnicity, rurality, and satisfaction with the psychological care received. Results show that (1) for each extra point in baseline depressive symptomatology (according to BDI scores), a reduction of 0.5 points in depressive symptomatology is expected at 12 weeks, keeping the value of the other variables constant and (2) for each additional point in satisfaction with the psychological care received, a reduction of 4.3 points in depressive symptomatology is expected at 12 weeks, keeping the value of the other variables constant. The interaction between the groups to which participants were randomized and their satisfaction with the care received was not significant (Table 4).

Discussion

Principal Findings

This clinical trial, conducted in 16 primary care centers located in vulnerable areas, far from the Chile's capital city, achieved recruitment rates of 60% of the initially projected sample and high response rates at 12-week follow-up (90.9%). The RCDC intervention significantly surpassed EUC in terms of user satisfaction with psychological treatment. Phone monitoring, a component of the intervention, was well received by the adolescents, and the primary care teams perceived that the intervention was useful and provided them with innovative tools for the comprehensive management of adolescent depression.

However, over one-third of the clinicians of the primary care teams had little or no participation on the Internet platform, and only one-third of the patients displayed an adequate adherence to the pharmacological treatment. Clinicians who participated in the RCDC group identified some obstacles in primary mental health care that may compromise the integrity of the intervention. Finally, at 12-week follow-up, RCDC for the treatment of adolescent depression did appear to have equivalent effectiveness compared with EUC, achieving comparable levels of depressive symptoms and HRQoL. Overall, more severe cases seemed to benefit more from treatment, as well as those who felt more satisfied with the psychological care received.

Strengths of the Study

This is the first report of a cluster randomized trial in a Latin American country that tests the feasibility, acceptability, and effectiveness of an ICT-assisted collaborative depression care intervention for the management of depressed adolescents from vulnerable backgrounds in remote areas lacking specialized mental health services.

Being able to train primary care teams and remotely implementing a complex intervention involving multiple components (RCDC) in several health care centers in a highly vulnerable area constitutes a major achievement, especially considering that the intervention was well received by the participants, increasing local teams' trust and fostering a positive perception of the adolescent depression treatment provided.

Therefore, this study contributes to scientific knowledge in multiple ways: not only does it respond to the marked lack of research equality in the field of depression in vulnerable populations [40], but it also addresses the sore need to generate mental health–related learning opportunities aimed at primary care professionals in Latin America and the Caribbean [41] and helps tackle the limited access to current scientific and technological advances available to health care teams and individuals from remote areas, thus expanding clinicians' problem-solving toolkit.

Limitations of the Study

However, this study has several design and implementation limitations that need be solved before conducting a larger clinical trial with suitable statistical power to draw conclusive results for the assessment of the effectiveness of the intervention.

Baseline Characteristics of the Primary Care Clinics Not Assessed

As the groups were not matched according to the characteristics of the participating primary care teams, there is a possibility that EUC practices performed better regarding depression management than RCDC practices before the project started. Additionally, the initial training of EUC clinicians might have improved their ability to treat depression. Both the characteristics of the participating health care providers and the robustness of the usual care provided have been identified to play a role in the effects of similar interventions [15,42].

Inadequate Statistical Power

Only 60% of the initially projected sample was recruited. The most likely explanation of this is the high turnover of clinicians in remote areas and the service pressure in usual practice, which limited the health care teams' ability to identify mental health problems in the population that they serve. To solve this issue, the researchers tried several strategies, such as expanding the recruitment period from 13 to 22 months, carefully reviewing the processes for detecting depressed adolescents in health care centers to strengthen their comprehensive assessment (eg, in sexual and reproductive health checkups conducted by midwives), establishing contact with nearby schools, and training personnel to detect and refer potentially depressed students.

Integrating Innovations Into Real-World Primary Care Settings

The use of the SEHR was a highly innovative tool aimed at improving the communication between health teams; however, it was not integrated into the daily routine of the participating physicians, and a limited time scheduled for this activity had to be negotiated with the local authorities. Moreover, the primary care teams were working in a context of high patient demand and regular staff turnover. Thus, registering patient information in the SEHR may have been experienced as burdensome by some of the physicians, partly explaining the low use of this system.

Over the course of the study, the researchers tried to complement the SEHR with more frequent in-person assistance to strengthen the bond and increase coordination between the local teams and specialists. However, this trial aimed to be pragmatic and to simulate real-life conditions rather than implement a highly complex, multifaceted intervention with high involvement of qualified study personnel, as those conducted in the United States [16,17]. In future implementations, if the integration of the SEHR into the currently working information system is not possible, it will be necessary, at least, to define protected time for using the Internet platform, explore elements that could make it more user-friendly, and possibly incorporate videoconferencing functionalities, depending on resource availability.

Patients' Preferences and Staff Stability

The low percentage of patients who displayed satisfactory adherence to the pharmacological treatment prescribed may compromise the integrity of the intervention, although this issue must be contextualized. Studies conducted in other depressed populations in Chile highlight the role of beliefs and attitudes toward medication in treatment adherence[43], and it is plausible that adolescents coming from highly vulnerable areas—as those participating in this study—are specially prone to noncompliance because of a combination of biological, psychological, and social factors.[44]. Moreover, the high turnover of clinicians in remote areas, especially physicians, along with the difficulties of training the incoming health care personnel could have resulted in a level of antidepressant prescription higher than that recommended by physicians without suitable training for managing adolescent depression.

It is noteworthy that one of the differences between the groups in terms of user satisfaction concerned psychological treatment and that this was a significant predictor for a decrease in depressive symptoms at 12-week follow-up in both groups. This component of both interventions was provided by the most stable personnel in Chilean primary care teams—psychologists. Therefore, psychologists could play a more relevant role in future interventions of this type. However, it is worthwhile to mention that as no other treatment process outcome was assessed in this trial, it is not clear what may have been the cause(s) of these differences favoring the RCDC intervention in terms of satisfaction and why this apparent advantage did not translate into differences in depressive symptoms across groups. As such, future studies of this type must integrate such assessments.

Other Issues

This study used a short follow-up period at the end of acute phase of treatment according to national guidelines [45], whereas studies employing longer intervention and follow-up periods have reported that collaborative care for depressed adolescents is effective [16,17,19]. Finally, it is known that depression management in primary care settings is suboptimal [11], and RCDC was further hampered by barriers related to patient (low consultation and adherence rates), provider (lack of time/interest to participate), and system issues (high clinician turnover). Therefore, the RCDC intervention may have lacked intensity and or or duration to cause the desired effect. Finally, it seems the RCDC group had a higher proportion of adolescents from more vulnerable backgrounds, with significant differences in socioeconomic status but also observable differences in terms of ethnicity and rurality. The intervention has not been adapted to treat this vulnerable population, and although evidence shows that school-aged children from minority ethnic groups may be at higher risk for depression [46], collaborative depression care programs have been shown to reduce racial disparities in adults [47].

Implications for Practice and Research

The implementation of this clinical trial in primary care centers in a vulnerable region of Chile was hindered by some structural characteristics of health services in the country. The high turnover level observed in health care teams, particularly involving physicians, prompts the need to develop a continuous distance training system that enables health care professionals who have recently joined a center to manage depressed adolescents. Likewise, the greater stability of psychologists in remote health care teams could make it possible to provide them with more intensive training in effective psychotherapy approaches for treating adolescent depression. Also, given their continuity, future initiatives could explore their role as case managers in charge of conducting a more thorough patient follow-up.

On the other hand, one of the main complaints received regarding the intervention concerned the user-friendliness of the Internet platform, the lack of time available to input case data and coordinate the teams, and the need to have specialists working in person with the primary care teams. In this respect, many primary care centers have electronic patient records; therefore, if specialists could establish contact using this medium, the intervention would become more user-friendly while also saving the local teams some time.

Research on collaborative care for the management of depression in adolescents is relatively scarce and limited to the United States. This clinical trial, carried out in a Latin American country, shows that these programs can be implemented in resource-poor settings and with vulnerable populations, reaching adequate acceptability levels among health care teams and patients. However, this study also highlights the difficulties involved in carrying out research in remote areas where the service pressure is high. The low levels of identification of depression conspired against achieving the targeted sample size.

Conclusions

It is feasible to implement a remote collaborative program for the management of depression among vulnerable adolescents from areas with poor access to specialized mental health services. The intervention was shown to be acceptable for the participants and showed a promising way to reduce the mental health care gap in remote areas. Before conducting a future clinical trial with enough statistical power to assess the program's effects on users' depressive symptoms and HRQoL, it is necessary to improve recruitment strategies, promote more interaction between local teams and specialists through the Internet platform, and test innovative strategies to prevent staff turnover in remote areas from affecting the abilities and joint work of health care teams. Nonetheless, the effectiveness of the intervention seems to be comparable with an already nationwide established treatment program that proved to be highly efficacious under controlled conditions [22,23].

Acknowledgments

This project was funded by the National Fund for Research and Development in Health Care (Fondo Nacional de Investigación y Desarrollo en Salud, FONIS), part of CONICYT and the Chilean Ministry of Health (Project SA 12I2294), and the Innovation Fund for Competitiveness (Fondo de Innovación para la Competitividad, FIC), part of the Ministry of Economy, Development, and Tourism, through the Millennium Scientific Initiative (Project IS130005). The authors wish to thank all participating primary care centers and users.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 532KB - jmir_v20i1e38_app1.pdf]

References

1. Wright DR, Katon WJ, Ludman E, McCauley E, Oliver M, Lindenbaum J, et al. Association of adolescent depressive symptoms with health care utilization and payer-incurred expenditures. Acad Pediatr 2016 Jan;16(1):82-89 [FREE Full text] [doi: 10.1016/j.acap.2015.08.013] [Medline: 26456002]
2. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Washington, DC: American Psychiatric Association; 2013.
3. Strandheim A, Bjerkeset O, Gunnell D, Bjørnelv S, Holmen TL, Bentzen N. Risk factors for suicidal thoughts in adolescence--a prospective cohort study: the Young-HUNT study. Br Med J Open 2014 Aug 20;4(8):e005867 [FREE Full text] [doi: 10.1136/bmjopen-2014-005867] [Medline: 25142264]
4. McManama O'Brien KH, Becker SJ, Spirito A, Simon V, Prinstein MJ. Differentiating adolescent suicide attempters from ideators: examining the interaction between depression severity and alcohol use. Suicide Life Threat Behav 2014 Feb;44(1):23-33 [FREE Full text] [doi: 10.1111/sltb.12050] [Medline: 23889515]
5. Conway KP, Swendsen J, Husky MM, He J, Merikangas KR. Association of lifetime mental disorders and subsequent alcohol and illicit drug use: results from the National Comorbidity Survey-Adolescent Supplement. J Am Acad Child Adolesc Psychiatry 2016 Apr;55(4):280-288. [doi: 10.1016/j.jaac.2016.01.006] [Medline: 27015718]
6. McLeod GF, Horwood LJ, Fergusson DM. Adolescent depression, adult mental health and psychosocial outcomes at 30 and 35 years. Psychol Med 2016 May;46(7):1401-1412. [doi: 10.1017/S0033291715002950] [Medline: 26818194]
7. Gibb SJ, Fergusson DM, Horwood LJ. Burden of psychiatric disorder in young adulthood and life outcomes at age 30. Br J Psychiatry 2010 Aug;197(2):122-127 [FREE Full text] [doi: 10.1192/bjp.bp.109.076570] [Medline: 20679264]

8. Zhou X, Hetrick SE, Cuijpers P, Qin B, Barth J, Whittington CJ, et al. Comparative efficacy and acceptability of psychotherapies for depression in children and adolescents: a systematic review and network meta-analysis. World Psychiatry 2015 Jun;14(2):207-222 [FREE Full text] [doi: 10.1002/wps.20217] [Medline: 26043339]
9. Qin B, Zhang Y, Zhou X, Cheng P, Liu Y, Chen J, et al. Selective serotonin reuptake inhibitors versus tricyclic antidepressants in young patients: a meta-analysis of efficacy and acceptability. Clin Ther 2014 Jul 01;36(7):1087-1095.e4. [doi: 10.1016/j.clinthera.2014.06.001] [Medline: 24998011]
10. Cipriani A, Zhou X, Del Giovane C, Hetrick SE, Qin B, Whittington C, et al. Comparative efficacy and tolerability of antidepressants for major depressive disorder in children and adolescents: a network meta-analysis. Lancet 2016 Aug;388(10047):881-890. [doi: 10.1016/S0140-6736(16)30385-3]
11. Mojtabai R, Olfson M, Han B. National trends in the prevalence and treatment of depression in adolescents and young adults. Pediatrics 2016 Dec;138(6):pii: e20161878. [doi: 10.1542/peds.2016-1878] [Medline: 27940701]
12. Soria-Saucedo R, Walter HJ, Cabral H, England MJ, Kazis LE. Receipt of evidence-based pharmacotherapy and psychotherapy among children and adolescents with new diagnoses of depression. Psychiatr Serv 2016 Mar;67(3):316-323. [doi: 10.1176/appi.ps.201500090] [Medline: 26725295]
13. Campo JV, Bridge JA, Fontanella CA. Access to mental health services: implementing an integrated solution. J Am Med Assoc Pediatr 2015 Apr;169(4):299-300. [doi: 10.1001/jamapediatrics.2014.3558] [Medline: 25664846]
14. Psychiatry. Washington, DC: American Psychiatric Association and Academy of Psychosomatic Medicine; 2016. Dissemination of Integrated Care Within Adult Primary Care Settings: the Collaborative Care Model URL: https://www.psychiatry.org/psychiatrists/practice/professional-interests/integrated-care/get-trained/about-collaborative-care [accessed 2018-01-23] [WebCite Cache ID 6wgOQZfyd]
15. Asarnow JR, Rozenman M, Wiblin J, Zeltzer L. Integrated Medical-Behavioral Care Compared With Usual Primary Care for Child and Adolescent Behavioral Health: A Meta-analysis. JAMA Pediatr 2015 Oct;169(10):929-937. [doi: 10.1001/jamapediatrics.2015.1141] [Medline: 26259143]
16. Asarnow JR, Jaycox LH, Duan N, LaBorde AP, Rea MM, Murray P, et al. Effectiveness of a quality improvement intervention for adolescent depression in primary care clinics: a randomized controlled trial. J Am Med Assoc 2005 Jan 19;293(3):311-319. [doi: 10.1001/jama.293.3.311] [Medline: 15657324]
17. Richardson LP, Ludman E, McCauley E, Lindenbaum J, Larison C, Zhou C, et al. Collaborative care for adolescents with depression in primary care: a randomized clinical trial. J Am Med Assoc 2014 Aug 27;312(8):809-816 [FREE Full text] [doi: 10.1001/jama.2014.9259] [Medline: 25157724]
18. Clarke G, Debar L, Lynch F, Powell J, Gale J, O'Connor E, et al. A randomized effectiveness trial of brief cognitive-behavioral therapy for depressed adolescents receiving antidepressant medication. J Am Acad Child Adolesc Psychiatry 2005 Sep;44(9):888-898. [Medline: 16113617]
19. Richardson L, McCauley E, Katon W. Collaborative care for adolescent depression: a pilot study. Gen Hosp Psychiatry 2009 Jan;31(1):36-45 [FREE Full text] [doi: 10.1016/j.genhosppsych.2008.09.019] [Medline: 19134509]
20. Wright DR, Haaland WL, Ludman E, McCauley E, Lindenbaum J, Richardson LP. The costs and cost-effectiveness of collaborative care for adolescents with depression in primary care settings. J Am Med Assoc Pediatr 2016 Nov 01;170(11):1048-1054. [doi: 10.1001/jamapediatrics.2016.1721]
21. Organisation for Economic Development and Cooperation. In It together: Why Less Inequality Benefits All. Paris: OECD; 2015.
22. Araya R, Rojas G, Fritsch R, Gaete J, Rojas M, Simon G, et al. Treating depression in primary care in low-income women in Santiago, Chile: a randomised controlled trial. Lancet 2003 Mar;361(9362):995-1000. [doi: 10.1016/S0140-6736(03)12825-5]
23. Araya R, Flynn T, Rojas G, Fritsch R, Simon G. Cost-effectiveness of a primary care treatment program for depression in low-income women in Santiago, Chile. Am J Psychiatry 2006 Aug;163(8):1379-1387. [doi: 10.1176/ajp.2006.163.8.1379]
24. Vöhringer PA, Jimenez MI, Igor MA, Fores GA, Correa MO, Sullivan MC, et al. Detecting mood disorder in resource-limited primary care settings: comparison of a self-administered screening tool to general practitioner assessment. J Med Screen 2013 Sep;20(3):118-124. [doi: 10.1177/0969141313503954]
25. World Health Organization, Ministry of Health of Chile. WHO-AIMS Report on Mental Health System in Chile. Santiago: World Health Organization and the Ministry of Health Chile; 2014.
26. Huawei Technologies Co, Ltd. Huawei. 2017. Harnessing the power of connectivity URL: http://www.huawei.com/minisite/gci/files/gci_2017_whitepaper_en.pdf?v=20170421 [WebCite Cache ID 6qEtxP78t]
27. Fortney JC, Pyne JM, Turner EE, Farris KM, Normoyle TM, Avery MD, et al. Telepsychiatry integration of mental health services into rural primary care settings. Int Rev Psychiatry 2015 Dec;27(6):525-539. [doi: 10.3109/09540261.2015.1085838] [Medline: 26634618]
28. 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]
29. Fortney JC, Pyne JM, Edlund MJ, Williams DK, Robinson DE, Mittal D, et al. A randomized trial of telemedicine-based collaborative care for depression. J Gen Intern Med 2007 Aug;22(8):1086-1093 [FREE Full text] [doi: 10.1007/s11606-007-0201-9] [Medline: 17492326]

30. Fortney JC, Pyne JM, Mouden SB, Mittal D, Hudson TJ, Schroeder GW, et al. Practice-based versus telemedicine-based collaborative care for depression in rural federally qualified health centers: a pragmatic randomized comparative effectiveness trial. Am J Psychiatry 2013 Apr;170(4):414-425 [FREE Full text] [doi: 10.1176/appi.ajp.2012.12050696] [Medline: 23429924]
31. Ministerio de Desarrollo Social. Santiago, Chile: Ministerio de Desarrollo Social; 2016. Informe de desarrollo social URL: http://www.ministeriodesarrollosocial.gob.cl/pdf/upload/IDS2017.pdf [accessed 2018-01-23] [WebCite Cache ID 6wgVuWUSB]
32. Días S, Santibañez D, Cortés A, Raczynski G, Contreras N, Bozo N. Observatorio Niñez y Adolescencia. Santiago, Chile: Observatorio Niñez y Adolescencia; 2016. Infancia cuenta en Chile 2016: Cuarto Informe URL: http://www.xn--observatorioniez-kub.cl/wp-content/uploads/2016/12/Informe_Infancia_Cuenta2016.pdf
33. Sheehan D, Sheehan K, Shytle R, Janavs J, Bannon Y, Rogers JE, et al. Reliability and validity of the Mini International Neuropsychiatric Interview for Children and Adolescents (MINI-KID). J Clin Psychiatry 2010 Mar;71(3):313-326. [doi: 10.4088/JCP.09m05305whi] [Medline: 20331933]
34. Beck AT, Steer RA, Carbin MG. Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. Clin Psychol Rev 1988 Jan;8(1):77-100. [doi: 10.1016/0272-7358(88)90050-5]
35. Alvarado R, Vega J, Sanhueza G, Muñoz MG. Evaluación del programa para la detección, diagnóstico y tratamiento integral de la depresión en atención primaria, en Chile. Rev Panam Salud Publica 2005 Nov;18(4-5):278-286 [FREE Full text] [doi: 10.1590/S1020-49892005000900008]
36. KIDSCREEN Group Europe. The KIDSCREEN Questionnaires. Quality of Life Questionnaires for Children and Adolescents-Handbook. Lengerich, Germany: Papst Science Publisher; 2006.
37. Sepúlveda PR, Molina GT, Molina CR, Martínez NV, González AE, Montaño ER, et al. [Validation of an instrument to measure health-related quality of life in Chilean children and adolescents]. Rev Med Chil 2013 Oct;141(10):1283-1292 [FREE Full text] [doi: 10.4067/S0034-98872013001000007] [Medline: 24522356]
38. Fritsch R, Araya R, Solís J, Montt E, Pilowsky D, Rojas G. Ensayo clínico aleatorizado de farmacoterapia con monitorización telefónica para mejorar el tratamiento de la depresión en la atención primaria en Santiago, Chile. Rev Méd Chile 2007 May;135(5):587-595. [doi: 10.4067/S0034-98872007000500006]
39. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011 Dec;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
40. Perrino T, Beardslee W, Bernal G, Brincks A, Cruden G, Howe G, et al. Toward scientific equity for the prevention of depression and depressive symptoms in vulnerable youth. Prev Sci 2015 Jul;16(5):642-651 [FREE Full text] [doi: 10.1007/s11121-014-0518-7] [Medline: 25349137]
41. Pan American Health Organization. Washington, DC: Pan American Health Organization; 2013. WHO-AIMS: Report on mental health systems in Latin America and the Caribbean URL: http://www.paho.org/hq/index.php?option=com_docman&task=doc_view&gid=21325&Itemid=270&lang=en
42. Shirazi M, Lonka K, Parikh SV, Ristner G, Alaeddini F, Sadeghi M, et al. A tailored educational intervention improves doctor's performance in managing depression: a randomized controlled trial. J Eval Clin Pract 2013 Feb;19(1):16-24. [doi: 10.1111/j.1365-2753.2011.01761.x] [Medline: 21883718]
43. Rojas G, Santelices M, Martínez P, Tomicic A, Reinel M, Olhaberry M, et al. Barreras de acceso a tratamiento de la depresión posparto en Centros de Atención Primaria de la Región Metropolitana: un estudio cualitativo. Rev Méd Chile 2015 Apr;143(4):424-432. [doi: 10.4067/S0034-98872015000400002]
44. Pacheco P B, Aránguiz G C. Factores relacionados a la adherencia a tratamiento en adolescentes con depresión. Rev Chil Neuro-psiquiatr 2011 Mar;49(1):69-78. [doi: 10.4067/S0717-92272011000100009]
45. Ministerio de Salud de Chile. Normassalud. Santiago: MINSAL; 2013. Guía clínica para el tratamiento de adolescentes de 10 a 14 años con depresión. Serie Guías Clínicas No GES URL: https://www.normassalud.com/documentos/chile/guia_depresion_2013.pdf [accessed 2018-01-23] [WebCite Cache ID 6wgcbUsHn]
46. Stirling K, Toumbourou JW, Rowland B. Community factors influencing child and adolescent depression: A systematic review and meta-analysis. Aust N Z J Psychiatry 2015 Oct;49(10):869-886. [doi: 10.1177/0004867415603129] [Medline: 26416916]
47. Davis TD, Deen T, Bryant-Bedell K, Tate V, Fortney J. Does minority racial-ethnic status moderate outcomes of collaborative care for depression? Psychiatr Serv 2011 Nov;62(11):1282-1288. [doi: 10.1176/ps.62.11.pss6211_1282] [Medline: 22211206]

Abbreviations

BDI: Beck Depression Inventory
EUC: enhanced usual care
HRQoL: health-related quality of life
ICTs: information and communication technologies
MDD: major depressive disorder

RCDC: remote collaborative depression care
SEHR: shared electronic health record

Edited by G Eysenbach; submitted 28.06.17; peer-reviewed by TR Soron, P Yellowlees, E Ludman; comments to author 03.08.17; revised version received 21.08.17; accepted 11.10.17; published 31.01.18.

Please cite as:
Martínez V, Rojas G, Martínez P, Zitko P, Irarrázaval M, Luttges C, Araya R
Remote Collaborative Depression Care Program for Adolescents in Araucanía Region, Chile: Randomized Controlled Trial
J Med Internet Res 2018;20(1):e38
URL: https://www.jmir.org/2018/1/e38/
doi:10.2196/jmir.8021
PMID:29386172

Original Paper

Paper- or Web-Based Questionnaire Invitations as a Method for Data Collection: Cross-Sectional Comparative Study of Differences in Response Rate, Completeness of Data, and Financial Cost

Jonas Fynboe Ebert[1], MD; Linda Huibers[1], MD, PhD; Bo Christensen[2], MD, PhD; Morten Bondo Christensen[1], MD, PhD

[1]Department of Public Health, Research Unit for General Practice, Aarhus University, Aarhus, Denmark

[2]Department of Public Health, Section for General Medical Practice, Aarhus University, Aarhus, Denmark

Corresponding Author:
Jonas Fynboe Ebert, MD
Department of Public Health
Research Unit for General Practice
Aarhus University
Bartholins Allé 2
Aarhus, 8000
Denmark
Phone: 45 25309292
Email: jonasebert@ph.au.dk

Abstract

Background: Paper questionnaires have traditionally been the first choice for data collection in research. However, declining response rates over the past decade have increased the risk of selection bias in cross-sectional studies. The growing use of the Internet offers new ways of collecting data, but trials using Web-based questionnaires have so far seen mixed results. A secure, online digital mailbox (e-Boks) linked to a civil registration number became mandatory for all Danish citizens in 2014 (exemption granted only in extraordinary cases). Approximately 89% of the Danish population have a digital mailbox, which is used for correspondence with public authorities.

Objective: We aimed to compare response rates, completeness of data, and financial costs for different invitation methods: traditional surface mail and digital mail.

Methods: We designed a cross-sectional comparative study. An invitation to participate in a survey on help-seeking behavior in out-of-hours care was sent to two groups of randomly selected citizens from age groups 30-39 and 50-59 years and parents to those aged 0-4 years using either traditional surface mail (paper group) or digital mail sent to a secure online mailbox (digital group). Costs per respondent were measured by adding up all costs for handling, dispatch, printing, and work salary and then dividing the total figure by the number of respondents. Data completeness was assessed by comparing the number of missing values between the two methods. Socioeconomic variables (age, gender, family income, education duration, immigrant status, and job status) were compared both between respondents and nonrespondents and within these groups to evaluate the degree of selection bias.

Results: A total 3600 citizens were invited in each group; 1303 (36.29%) responded to the digital invitation and 1653 (45.99%) to the paper invitation (difference 9.66%, 95% CI 7.40-11.92). The costs were €1.51 per respondent for the digital group and €15.67 for paper group respondents. Paper questionnaires generally had more missing values; this was significant in five of 17 variables ($P<.05$). Substantial differences were found in the socioeconomic variables between respondents and nonrespondents, whereas only minor differences were seen within the groups of respondents and nonrespondents.

Conclusions: Although we found lower response rates for Web-based invitations, this solution was more cost-effective (by a factor of 10) and had slightly lower numbers of missing values than questionnaires sent with paper invitations. Analyses of socioeconomic variables showed almost no difference between nonrespondents in both groups, which could imply that the lower response rate in the digital group does not necessarily increase the level of selection bias. Invitations to questionnaire studies via digital mail may be an excellent option for collecting research data in the future. This study may serve as the foundational pillar of digital data collection in health care research in Scandinavia and other countries considering implementing similar systems.

(J Med Internet Res 2018;20(1):e24) doi:10.2196/jmir.8353

KEYWORDS

questionnaire study; response rate; completeness of data; financial costs; missing values; selection bias; digital post; digital survey invitation; Web-based questionnaire

Introduction

The preferred mode for collecting survey data in research has traditionally been the paper questionnaire [1], which is a simple and palpable way of communicating between citizen and researcher. However, in recent years, this way of collecting data has been challenged. Over the last decade, response rates have declined by approximately 1% per year in many countries [1-4]. A low response rate may induce selection bias because respondents may differ systematically from nonrespondents, and the study population will thus not represent the target population [5]. The costs of sending letters by surface mail in the Scandinavian countries have also increased markedly in the last years. For example, a cost increase of 90% was seen in Denmark in 2016 [6,7]. Additionally, longer delivery time (up to 8 days) and fewer post offices may also imply that many Danes now tend to check their physical mailbox less often (M Christensen, email communication, June 5, 2017 and [7,8]).

Data collection by paper questionnaire with traditional surface mail involves several time-consuming and costly steps, such as printing and packing questionnaires and scanning returned questionnaires. A study performed in Denmark in 2013 estimated an average expenditure of €8.62 per respondent, exclusive of motivational costs and researcher time [1]. Filling out a paper questionnaire is a practical and easy method because the respondent only needs a pen and time to participate. Yet, this option also enables the respondent to leave questions unanswered (intentionally or unintentionally), to fill in more answers for one question than allowed, or to check a box outside the intended boundaries. These errors often result in missing values, especially when completed questionnaires are read by a machine, and compromise the data.

The growing use of the Internet has made the Web-based questionnaire an obvious alternative to the paper questionnaire. Web-based studies have been shown to lower the data collection costs [9,10], which is attractive, especially in large population-based surveys [1]. As more and more people have access to the Internet, this has reduced potential variations in the population coverage between paper- and Web-based questionnaires and lowered the risk of selection bias from using the Internet for questionnaire surveys [4]. In 2016, 94% of all Danish citizens had access to a computer and the Internet at home [11]. Furthermore, free public access to the Internet is offered at all Danish libraries, which ensures almost 100% access to the Internet for the entire population.

The Danish government made it mandatory in 2014 for all Danish citizens older than 15 years of age to use a digital mailbox when communicating with public authorities [12]. In exceptional cases, citizens may opt out of this arrangement and receive information by surface mail. The standard arrangement with digital post has two forms: a personal mailbox for each citizen, which is accessed at the public website borger.dk [13], and an electronic mailbox, which is managed by the private company E-boks [14]. The Danish public authorities have thus successfully ensured that the vast majority (89.3%) of the Danish population have access to a secure and low-cost way of receiving letters from many different public authorities. This includes pay slips, invoices, and notices from the national health care services [15].

The availability of digital mail for the majority of Danish citizens makes it possible to study the use of Web-based invitations to scientific studies. To our knowledge, no former study has investigated the impact of being able to reach such a large percentage of the public in terms of data collection.

In this study, we aim to compare the response rates for invitations to a questionnaire-based study sent out by traditional mail and by digital mail. We also aim to compare the completeness of data and the costs of these two ways of collecting data for research.

Methods

Design

We conducted a cross-sectional comparative study using data obtained from a previous study, which was performed by one of the authors, and a new data collection, which we performed 8 months later. The previous study was based on invitations sent by traditional surface mail (paper group). Respondents could complete an enclosed paper questionnaire or an online questionnaire accessed through a 12-digit code. The new data collection used an invitation sent to an electronic mailbox. This invitation included a unique Web link to a Web-based version of the same questionnaire (digital group). Citizens without an electronic mailbox received the invitation and a paper questionnaire by surface mail. Figure 1 outlines the main components of the two different invitation methods.

The Questionnaire

The questionnaire had already been developed; it originally formed part of two yet-unpublished studies on citizen help-seeking behavior by Huibers et al (unpublished data, 2017) and Keizer et al (unpublished data, 2017). The aim of this study was to investigate the help-seeking behavior among citizens in need of acute health care during out-of-office hours and the factors related to frequent requests of out-of-hours health care. The questionnaire was sent to parents of children aged 0 to 4 years, to citizens aged 30 to 39 years, and to citizens aged 50 to 59 years. The main part of the questionnaire consisted of six cases presenting well-defined acute health problems; each case was followed by a question on help-seeking behavior with nine multiple response options (including “other”). The parents of young children were presented with different cases than the adult citizens.

Figure 1. Flowchart of invitation methods and distribution of response methods. Percentages in brackets describe the distribution in the study.

Furthermore, 21 questions focused on factors related to help-seeking behavior and background characteristics; these included items from previously validated scales and self-developed questions. The questionnaire regarding children had four additional questions. Questions about factors related to help seeking often included several response options (up to 10 per question) measured by a Likert scale (from a four-point and up to a seven-point scale). Mostly, only one answer was possible. For some questions, the response category "don't know" and/or "not relevant" was also an option. The background characteristics included a question on age (for which a number should be stated) and several questions with the response option "other" including a free-text field for explanatory comments. Finally, the questionnaire had an extra free-text box for additional comments.

Results from the original study have not yet been published. To our knowledge, this questionnaire is fairly representative of questionnaires used in this field of research. The Web-based questionnaire did not require every question to be answered; this design was chosen to ensure that the method was comparable to the paper questionnaire.

Inclusion and Exclusion Criteria

In the original study by Huibers et al (unpublished data, 2017), three age groups including 1200 people each were available to compare the two ways of sending out questionnaires: 0 to 4 years (parents), 30 to 39 years, and 50 to 59 years. To get a realistic impression of the all-round response rate, a fourth age group of 70 to 79 years was added. However, this age group was not included in the comparisons of costs and missing values because no information on paper-based data collection was available for this group. We also chose to include 1200 people in each age group because this allowed us to detect a minimum difference in the response rate of 3.8% between the paper group and the digital group if using a significance level of .05 and a power of .9. All citizens included in our study were randomly selected by the Statens Serum Institut, the Danish national institute for health data and disease control, using sex, age, and region to obtain a sample that was representative of the entire country.

The following groups were excluded: citizens living in institutions, citizens with publicly recorded protection against participating in research, and deceased citizens. Furthermore, citizens and siblings of children aged 0 to 4 years who had participated in the paper questionnaire study were excluded from the digital study. Parents of children aged 0 to 4 years were also excluded from the other age groups to ensure that they did not receive two questionnaires. Citizens who returned an empty questionnaire were excluded because of suspicion of disease (eg, dementia or autism) or insufficient language skills. Likewise, citizens who stated that they wished not to participate were registered as nonrespondents. This group of excluded citizens accounted for 0.7% in the paper study and 0.9% in the digital study.

Data Collection

A total of 89.3% of the Danish population older than 15 years had access to their individual digital mailbox in June 2016. A recent opinion poll showed that 90% of all users reported to read their digital mailbox “always or often” [15,16]. Citizens are linked with the digital mailbox through their unique civil registration number [14]. Citizens younger than 15 years automatically belong under their parents’ digital mailbox; this means that either the mother or both parents receive their digital mail. Citizens can apply for exemption from the digital mailbox in case of cognitive or physical disability, no access to a computer with Internet, language barriers, or if not living in Denmark.

A paper questionnaire containing 27 to 31 questions and subcategories, depending on age group, was sent to the paper group in November 2015 using the traditional method of mailing paper questionnaires (ie, surface mail including a stamped return envelope). After 21 days, a reminder (including a questionnaire and stamped return envelope) was sent by surface mail to nonrespondents.

A similar email invitation letter containing a personalized active link to a Web-based version of the questionnaire (using Survey Xact) was sent to the digital mailbox of the digital group in June 2016. The email had a personalized active link that directed the respondent directly to the questionnaire without using any key or password. Nonrespondents received a first reminder after 7 days and a second reminder after 14 days, both using the digital mailbox. The 3.6% of the three youngest age groups (0-4, 30-39, and 50-59 years) in the digital group and the 30.4% of the oldest age group (70-79 years) who did not have a digital mailbox received a paper questionnaire. If they did not respond within 21 days, a reminder was sent by surface mail (see Figure 2). The wording of the invitation was the same for both digital and paper groups, except for the explanation of how to access the Web-based questionnaire. Both groups had the same incentive: to enter a draw to win two cinema tickets.

After the data collection, we received additional data from Statistics Denmark on all citizens included in the study (N=8382). These additional data included family income, immigrant status (born in Denmark or immigrant), education, and job status (employed or unemployed). This information was used to compare respondents and nonrespondents and to compare nonrespondents in the digital group to nonrespondents in the paper group. These comparisons enabled us to assess the level of selection bias. The hypothesis was that no difference between the two groups of nonrespondents would mean that the same degree of selection bias was present in the two methods.

Outcome Variables

The following primary outcome measures were investigated: difference in response rates, completeness of data, and financial costs. A response was considered valid when a questionnaire was returned by mail or completed online. If a returned questionnaire was blank, it did not count as a response (and was excluded). The completeness of data was assessed by measuring the percentage of missing values in the 17 variables that were considered the most important by the author team who conducted the questionnaire study. To calculate the costs, we measured the time spent on distributing the questionnaires by timing every aspect of the distribution and the data collection for both paper- and Web-based questionnaires. Every expense in the process was registered and summed to calculate a mean cost per respondent.

Statistical Analyses

When comparing the two distribution groups and respondents to nonrespondents, we used Student *t* test for the variables age and family income. The two-sample test for proportions was used when assessing gender, immigrant status, and job status, and the chi-square test was used to compare educational status. Comparisons were made between and within response groups.

Figure 2. Flowchart of participants and percentage with digital mailbox. The bottom box states the percentage of the age group with access to the digital mailbox (E-boks).

Differences in response rates, overall and within each age strata, were estimated using generalized linear models with ID link and the binomial family. The completeness of data was assessed by calculating the percentages of missing values in the 17 variables that were found most relevant for the study (as described previously) using the two-sample *z* test for proportions. The difference in total costs between the paper group and the digital group was also estimated using generalized linear models with ID link and the binomial family. Only age groups 0 to 4, 30 to 39, and 50 to 59 years were compared in the tables.

We used Stata 14 (StataCorp, College Station, TX, USA) for the statistical analyses.

Ethical Considerations

The project was approved by the Danish Data Protection Agency (j no 2015-57-0002, AU j no 62908 218). According to Danish law, approval from the Committee on Health Research Ethics of the Central Denmark Region was not needed because the study did not include biomedical intervention.

Results

The total study population consisted of 8400 people. However, a total of 18 people were excluded: 10 were already included in the paper group, two had an already included sibling, and six had died. More than 96.38% (3458/3588) of the sample had access to a digital mailbox, except in the age group 70 to 79 years in which only 69.58% (835/1200) had access (Figure 2).

Respondents Versus Nonrespondents

Only a few significant differences were found between the paper group and the digital group within respondents and nonrespondents with respect to age, gender, family income, immigrant status, education and job status in all age groups. Exceptions were job status for respondents in age group 50 to 59 years, family income in age groups 0 to 4 years and 50 to 59 years, age for nonrespondents in age group 30 to 39 years and for respondents in age group 0 to 4 years, and education in age group 50 to 59 years, where a significant difference was found between the digital group and the paper group (*P*<.05) (Table 1).

When comparing respondents to nonrespondents, we saw significant differences (*P*<.05) in age (0-4 years), family income, immigrant status, job status, and education length of more than 15 years (Table 1). In the education variable, 597 missing values were generated; 459 of these concerned immigrants.

Nonrespondents tended to have lower income, to have shorter education, and to be less likely to be employed than respondents. Additionally, more nonrespondents were male.

Response Rate

The overall response rate in the digital group was 36.31% (1303/3588), almost 10 percentage points lower than in the paper group (45.99%, 1653/3594) (Table 2). In every age group, the response rate was lower in the digital group than in the paper group; the largest difference was seen in the age group 30 to 39 years (paper: 35.81%, 429/1198; digital: 23.18%, 277/1195). In the paper group, 334 of 1653 (20.21%) answered online. In the digital group, 1280 of 1303 (98.23%) answered online.

For the age group 70 to 79 years, the response rate was 50.58% (607/1200). For this age group, no significant difference in response rate was seen between the group who received the invitation digitally (69.58%, 835/1200) and the group who received it on paper (30.42%, 365/1200; difference 1.0%, 95% CI –9.6% to 7.7%).

Financial Costs

The total costs of collecting data in the digital group was €1969.16 (Table 3). This figure was considerably lower than in the paper group, which amounted to €25,905.28. Although the response rate was lower in the digital group, the costs per respondent were markedly lower than in the paper group (€1.51 vs €15.67). We found that costs related to wages for handling of the digital questionnaires were less than half (43.40%, €926/€2133) of the costs related to handling of paper questionnaires (Multimedia Appendix 1). This lower figure was found even though much more time was spent on handling paper questionnaires, but this work was conducted by assistants who were remunerated at lower wages than the researchers were. We also found that much more time was spent on handling paper questionnaires than on the digital questionnaires (118 hours vs 39 hours, *P*<.001).

Completeness of Data

Table 4 shows the percentage of missing values in the variables that were found most important by the author team conducting the questionnaire study. The number of missing values was generally lower in the digital group; this was significant in five of 17 variables on a total of nine of 68 occasions.

Table 1. Background data for respondents versus nonrespondents, including age, gender, family income, immigrant status, education, and job status (N=7182).

Background data	Invitation method and age range (years)					
	Digital			Paper		
	0-4	30-39	50-59	0-4	30-39	50-59
Respondents						
n	495	277	531	572	429	652
Age (years), mean (95% CI)	33.3[a] (32.9, 33.7)	34.5 (34.2, 34.9)	54.5 (54.2, 54.7)	34[a,b] (33.6, 34.4)	34.7 (34.4, 35.0)	54.2 (54.0, 54.5)
Gender (male), % (95% CI)	0.2[b] (–0.2, 0.6)	39.7[a] (33.9, 45.5)	43.9[a] (39.6, 48.1)	50.4[b] (46.2, 54.5)	37.9[a] (33.3, 42.5)	45.1[a] (41.3, 49.0)
Family income (€), mean (95% CI)	34,193 (32,863, 35,523)	35,804[a] (34,183, 37,424)	46,368[a] (43,892, 48,844)	35,546[a] (34,426, 36,666)	35,280[a] (33,066, 37,494)	44,467[a] (42,022, 46,912)
Immigrant status (Danish), % (95% CI)	92.2 (89.8, 94.5)	89.5 (85.9, 93.1)	95.9 (94.2, 97.6)	90.0[a] (87.5, 92.4)	89.9[a] (87.1, 92.8)	95.0[a] (92.7, 96.2)
Education (years), % (95% CI)						
<10	7.3[a] (5.2, 9.9)	8.7[a] (5.8, 12.7)	16.8[a],[b] (13.8, 20.3)	8.1[a] (6.1, 10.7)	8.5[a] (6.2, 11.7)	20.2[a,b] (17.3, 23.5)
10, 15	30.6[a] (26.7, 34.9)	33.1[a] (27.7, 39.0)	44.5[a,b] (40.2, 48.8)	30.4[a] (26.7, 34.4)	36.3[a] (31.8, 41.1)	49.0[a,b] (45.2, 52.8)
>15	62.1[a] (57.7, 66.3)	58.3[a] (52.2, 64.1)	38.7[a,b] (34.6, 43.0)	61.5[a] (57.3, 65.5)	55.1[a] (50.3, 59.9)	30.8[a,b] (27.4, 34.5)
Job status (employed), % (95% CI)	80.1[a] (76.6, 83.6)	84.1[a] (79.7, 88.4)	82.5[a,b] (79.2, 85.7)	78.3[a] (75.0, 81.8)	81.8[a] (78.2, 85.5)	87.5[a,b] (84.9, 90.0)
Nonrespondents						
n	698	918	669	626	769	546
Age (years), mean (95% CI)	32.2[a] (31.7, 32.6)	34.3 34.1, 34.5)	54.2 (54.0, 54.4)	32.3[a] (31.8, 32.7)	34.8[b] (34.6, 35.0)	54.2 (54.0, 54.5)
Gender (male), % (95% CI)	1.0[b] (0.2, 1.7)	54.0[a] (50.1, 57.3)	56.5 (52.7, 60.3)	51.7[b] (47.8, 55.6)	55.0[a] (51.5, 58.5)	54.6[a] (50.4, 58.8)
Family income (€), mean (95% CI)	32,275[b] (29,871, 34,678)	31,589[a] (29,737, 33,441)	40,800[a,b] (38,678, 42,922)	28,987[a,b] (27,883, 30,091)	30,481[a] (29,222, 31,740)	36,913[a,b] (35,232, 38,594)
Immigrant status (Danish), % (95% CI)	76.2[a] (73.0, 79.3)	78.5[a] (75.9, 81.2)	90.1[a] (87.8, 92.4)	79.4[a] (76.2, 82.6)	75.6[a] (75.6, 81.5)	87.2[a] (84.4, 90.0)
Education (years), % (95% CI)						
<10	17.6[a] (14.7, 20.8)	16.4[a] (14.0, 19.1)	26.3[a] (23.0, 29.8)	19.8[a] (16.7, 23.4)	21.1[a] (18.1, 24.4)	29.3[a] (25.5, 33.4)
10, 15	39.1[a] (35.3, 43.0)	44.4[a] (41.0, 47.9)	44.8[a] (41.0, 48.7)	38.2[a] (34.1, 42.3)	43.4[a] (39.7, 47.3)	45.0[a] (40.7, 49.3)
>15	43.4[a] (39.5, 47.3)	39.2[a] (35.8, 42.6)	28.9[a] (25.5, 32.5)	42.0[a] (39.9, 46.3)	35.5[a] (31.9, 39.2)	25.7[a] (22.1, 29.7)
Job status (employed), % (95% CI)	65.2[a] (61.7, 68.8)	77.2[a] (73.7, 77.8)	76.7[a] (73.5, 79.9)	63.4[a] (59.6, 67.2)	73.9[a] (70.8, 77.1)	71.8[a] (68.0, 75.7)

[a] Significant difference (*P*<.05) between respondents and nonrespondents in the same age group and distribution group.

[b] Significant difference (*P*<.05) between paper group and digital group in the same age group and response group. In the age group 0 to 4 years for the paper group, the mean age was calculated on the basis of the mother to ensure compatibility with the digital group. The percent of males in the gender variable for age group 0 to 4 years was low in the digital group because invitations were sent only to the mother, except in cases where the father had sole custody. The paper invitation was sent in the child's own name directly to the child's registered postal address.

Table 2. Response rates in different age groups for the two ways of collecting data.

Age group[a]	Paper group, n/sent (%)	Digital group, n/sent (%)	Difference, % (95% CI)
0-4 years	572/1198 (47.75)	495/1193 (41.49)	6.25 (2.28-10.23)
30-39 years	429/1198 (35.81)	277/1195 (23.18)	12.63 (9.01-16.25)
50-59 years	652/1198 (54.42)	531/1200 (44.25)	10.17 (6.19-14.16)
All	1653 (45.99)	1303 (36.32)	9.68 (7.41-11.94)

[a] Age group 70 to 79 years was not included in the final response rate.

Table 3. Costs (in €) for the two ways of collecting data[a].

Subject	Paper group (n=1653)	Digital group (n=1303)
Paper questionnaire and envelope	9892	383
Postage	13,815	479
Packaging and registration	768	21
Scanning and coding	1365	17
Digital postage	—	529
Coding	—	394
Handling of digital distribution	—	81
Incentives (draw for 2×2 tickets)	65	65
Total costs	25,905	1969
Costs per respondent	15.67	1.51

[a] Age group 70 to 79 years was not included in this analysis. For further details on costs, see Multimedia Appendix 1.

Table 4. Overview of number of missing values in percentage of responses for the two ways of collecting data for different age groups.

Variable	Digital group (years), % missing values				Paper group (years), % missing values			
	0-4 (n=495)	30-39 (n=277)	50-59 (n=531)	Total (n=1303)	0-4 (n=572)	30-39 (n=429)	50-59 (n=652)	Total (n=1653)
Background characteristics								
Age	1.01	0.36	0.56	0.69	1.40	1.40	1.07	1.27
Gender	1.62	1.08	1.51	1.46	1.75	1.17	1.38	1.45
Married/cohabiting	1.62	0.72	0.75	1.07	1.75	2.56	2.15[a]	2.12[a]
Education	0.20	0.72	0.75	0.54	0.70	0.93	0.61	0.73
Job	0.40	0.36	0.75	0.53	1.22	0.93	0.92	1.03
Ethnicity	0.81	1.44	3.20	1.99	12.45[a]	1.86	11.99	2.12
Factors related to help seeking outside office hours								
Choice to contact out-of-hours care	5.86	8.66	13.60	9.59	4.20	7.23	11.04	7.68
Self-efficacy	0.20	1.08	0.56	0.53	0.87	1.40	1.69	1.33[a]
Anxiety	1.62	0.36	2.26	1.61	1.05	0.93	1.38	1.15
Social support	0.20	0.36	0.38	0.31	0.87	1.40	1.84[a]	1.39[b]
Health literacy, navigation	0.40	0.72	1.69	0.99	0.70	0.93	2.15	1.33
Health literacy, information	0.20	0.00	0.75	0.38	0.52	1.17	1.38	1.03[a]
Right/barrier	1.21	2.89	2.82	2.22	1.40	2.56	3.53	2.54
Frequency	0.81	1.81	3.58	2.15	0.87	1.63	1.38[a]	1.27
Satisfaction with general practitioner	0.40	0.72	1.32	0.84	1.22	0.93	1.23	1.15
Satisfaction with out-of-hours care	1.41	0.72	2.07	1.53	1.57	1.63	1.84	1.69
Travel time	0.00	0.00	0.56	0.23	1.05[a]	0.70	1.23	1.03[b]

[a] Statistically significantly more missing values than in the digital group (*P*<.05).

[b] Statistically significantly more missing values than in the digital group (*P*<.01).

Discussion

Main Findings

In this questionnaire study, we obtained a significantly lower response rate for invitations sent out by a mandatory secure digital mailbox (36.32%, 1303/3588) than for paper invitations combined with paper questionnaires sent out by surface mail (45.99%, 1653/3594). This difference was seen for all three youngest age groups: parents of children aged 0 to 4 years, citizens aged 30 to 39 years, and citizens aged 50 to 59 years. Citizens aged 70 to 79 years had a high response rate (50.58%, 607/1200).

When exploring the completeness of data, we found that the paper questionnaires generally had more missing values; this was significant in five of 17 variables although variations were found for different age groups.

The costs were markedly lower for the digital method: €1.51 for the digital mailbox versus €15.67 for the paper questionnaire when calculated per respondent.

Strengths and Limitations

We were able to include a large group of Danish citizens with a secure digital mailbox. This provided us with good power for our analyses, which was a major strength of the study. In addition, we were able to compare nonrespondents in both groups (ie, digital invitation and paper invitation) on socioeconomic variables, which enabled us to evaluate potential selection bias related to applying two different data collection methods. By choosing four age groups that combined represented a substantial part of the population, we were able to explore the general applicability of a digital solution. Still, it is unknown if the response rates for other age groups would be similar to the ones we found.

It was a limitation that the two questionnaires were not sent out at the same time of the year (paper version in November 2015 and digital version in June 2016). Hence, we cannot rule out that some seasonal variation may have occurred. On the one hand, more people are usually ill in November/December, which may have lowered the likelihood of participation and thus lead to an underestimation of the difference in response rates [17]. On the other hand, as mentioned in the Introduction, the general decrease in response rates could have pulled in the other direction. However, we expect this potential variation to be insignificant. In addition, we received some complaints from citizens regarding technical difficulties with the active link in the invitation; this could have had a negative influence on the digital response rates and thus have lead to an overestimation of the difference in response rates.

The questionnaire was not linked to a specific contact with the health care services, which might also have resulted in lower response rates. This also implied that we were unable to make any assumptions about response rates, costs, and completeness of data for specific contacts. Furthermore, we aimed to compare traditional paper invitations and questionnaires with Web-based invitations and questionnaires. However, when sending out paper invitations, it is now common practice to include an option to go online and answer a Web-based version of the questionnaire [1,4]. In our study, 20.21% (334/1653) of the respondents from the paper group used the Web-based option. We chose to include them in the paper group as this realistically reflected the data collection method, but this might also have resulted in lower percentages of missing values and lower overall costs in the paper group compared to complete use of paper questionnaires exclusively. Furthermore, in the calculation of response rates, we also included the small fraction of citizens (in age groups 0-4, 30-39, and 50-59 years) who did not have the digital mailbox and thus received paper questionnaires. The reason was that using only digital invitations for our calculations would not provide us with a precise estimate of the response rate, whereas including the small fraction with no digital mailbox offered a more realistic reflection of the data collection method. The response rate would have been 37% (one percent point higher) if we had included only the digital invitations.

Interpretation of Results

Sending out paper questionnaires is a costly and time-consuming process [1,4]. Several studies have compared sending out invitations to paper-based and Web-based questionnaires. The overall trend shows little or no difference in the response rates between the two different data collection modes, with response rates ranging from 53% to 92% for the Web-based method and from 56% to 92% for the paper-based method [1,4,18-20]. However, our study found a significant difference in the response rates when comparing paper-based and Web-based methods (46%, 1653/3594 vs 36%, 1303/3588). The biggest difference in response rates was seen in the age group 30 to 39 years (12.7%), which is significantly higher (*P*<.001) than seen for parents of the age group 0 to 4 years (6.1%). An interesting finding was that the age group 30 to 39 years, with a mean age of 34.5 (SD 3.0) years, had significantly lower response rates than the parents of age group 0 to 4 years, with a mean age of 33.3 (SD 4.6) years (23.18%, 277/1195 vs 41.49%, 495/1193, *P*<.001). We believe that this was unrelated to sending out invitations only to the mothers of children aged 0 to 4 years in the digital group because approximately half of the respondents for children aged 0 to 4 years in the paper group were male. Studies using questionnaires concerning a specific contact generally get higher response rates [4]. This could be part of the explanation for the lower response rate in our study compared to the ones found in otherwise comparable studies [1,4,18-20].

Web-based questionnaires have repeatedly been proven to lower data collection costs [1,9,10]. Our study supports these findings as we saw a cost difference by a factor of 10 (€15.67 vs €1.51). In our calculations, we have taken into account the time it takes to handle paper and digital questionnaires. As the costs relating to wages in the handling of digital questionnaires are less than half (43%) the amount relating to the handling of paper questionnaires, this emphasizes that far more time is spent on handling paper questionnaires than digital questionnaires (118 hours vs 39 hours, *P*<.001). If we consider the completeness of the responses, earlier studies have shown that Web-based questionnaires have fewer missing values [21,22]. We looked at the percentage of missing values in a predefined number of variables, and our findings confirm that the completeness of data is higher when the invitation and response method is Web-based.

Van Gelder et al [4] stated in 2010, “Recent studies have already shown that respondents to Web-based questionnaires are comparable to those responding to traditional modes of data collection in terms of age, gender, income, education....” This is also in line with our findings. Furthermore, our analyses indicated that nonrespondents in the paper group and in the digital group did not differ from each other, except for family income, which implies that selection bias might not be a bigger problem in Web-based data collection than in traditional paper-based data collection [2,4].

Moreover, the older age groups tend to have less experience with using computers and the Internet [20], which is shown in the fact that only approximately 70% of this age group had access to a digital mailbox. However, in the near future, we expect to see an increasing use of the Internet in the older age groups, which could facilitate the use of electronic invitations and questionnaires and make it the new way of collecting data. In addition, we expect easier access to the digital mailbox, and the population will become more accustomed to checking their digital mailbox regularly, which might also help improve response rates in questionnaire studies.

Implications for Practice and Future Studies

Using digital mail as a new way of sending out questionnaires could be the future approach in questionnaire-based research. Because of the reduced costs, sending out more questionnaires for the same amount of money as that required for a paper-based trial could increase the power of a study. Potential selection bias remains an issue as in other questionnaire-based studies, although our results show that the direction of selection bias is similar for both methods of collecting data.

We chose not to include expenses for software to handle the Web-based questionnaires and hardware to handle the paper responses as this was available at our research unit as part of a university institution. Nevertheless, if an independent research group was to buy licenses to conduct Web-based surveys, these would entail considerable costs, which should be considered in the planning.

Citizens are increasingly flooded with online invitations to participate in evaluations and questionnaire surveys. It is important to develop standards when using digital mail for scientific research to maintain this method as a valid way of collecting data in the years to come. One option could be to use only digital mail when inviting citizens to a health-related questionnaire study regarding a specific contact with the health care system. It is our intention to test this in an upcoming study in the Danish out-of-hours primary care setting at the end of 2017 [23]. Future studies could compare response rates for different types of questionnaires when using this option. Another way of achieving acceptable response rates could be to use consumer panels, such as the global network operated by Kantar Active in 90 countries under different names (eg TNS Nipo in the Netherlands and TNS Gallup in Denmark), for which expenses saved are used to pay respondents [24].

Conclusion

A digital platform offering a secure communication system and targeting the individual citizen combined with easily accessible Web-based questionnaires to be completed on a mobile phone, tablet, or computer seems to be a low-cost option in future survey studies. Such a digital solution also appears to give higher completeness of data compared to data collection by paper invitations combined with paper questionnaires sent by surface mail. Lower response rates (especially in the younger age groups) could be a problem. Still, our findings suggest that the two methods exhibit similar degrees of selection bias for socioeconomic variables.

The secure digital post solution used in our study (e-Boks) is now available in Denmark, Norway, and Sweden, and the system has more than 12 million users in these countries. In the near future, we could see a massive shift from paper to digital data collection in questionnaire research. This project could serve as the foundational pillar of digital data collection and help us obtain a better understanding of the feasibility of this method in future health care research.

Acknowledgments

The author group would like to thank the respondents for taking the time to participate in this project. The author group also would like to acknowledge Hanne Beyer for database management and Lone Niedziella for proofreading the manuscript.

Authors' Contributions

JFE was the main author of this manuscript. MBC, LH, and BC contributed equally to supervision of the trial and review of the manuscript. LH sent out the paper invitation including the questionnaire as a part of her study of citizens’ behavior in the out-of-hours primary care setting in Denmark in 2015. JFE sent out the digital invitation in 2016.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Complete overview of financial costs.

[XLSX File (Microsoft Excel File), 18KB - jmir_v20i1e24_app1.xlsx]

References

1. 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]
2. Nohr EA, Frydenberg M, Henriksen TB, Olsen J. Does low participation in cohort studies induce bias? Epidemiology 2006 Jul;17(4):413-418. [doi: 10.1097/01.ede.0000220549.14177.60] [Medline: 16755269]
3. Morton LM, Cahill J, Hartge P. Reporting participation in epidemiologic studies: a survey of practice. Am J Epidemiol 2006 Feb 01;163(3):197-203. [doi: 10.1093/aje/kwj036] [Medline: 16339049]
4. van Gelder MM, Bretveld RW, Roeleveld N. Web-based questionnaires: the future in epidemiology? Am J Epidemiol 2010 Dec 1;172(11):1292-1298 [FREE Full text] [doi: 10.1093/aje/kwq291] [Medline: 20880962]
5. Griffin JM, Simon AB, Hulbert E, Stevenson J, Grill JP, Noorbaloochi S, et al. A comparison of small monetary incentives to convert survey non-respondents: a randomized control trial. BMC Med Res Methodol 2011 May 26;11:81 [FREE Full text] [doi: 10.1186/1471-2288-11-81] [Medline: 21615955]
6. Wichmann S. Berlingske Business. Posten varsler gigantisk portoforhøjelse [The Danish postal services announce giant price increases] URL: http://www.business.dk/transport/posten-varsler-gigantisk-portoforhoejelse [accessed 2017-05-04] [WebCite Cache ID 6qCo82yCn]
7. Schjølberg K. Frankering. Myten om portotakster [The myth about postage rates] URL: http://www.frankering.com/forum/topics/myten-om-portotakster [WebCite Cache ID 6qCoJFKme]
8. Henriksen M. DR. 2016 Apr 27. Leveringstid på breve vil stige voldsomt [Longer delivery times for Danish letters] URL: https://www.dr.dk/nyheder/politik/leveringstid-paa-breve-vil-stige-voldsomt [accessed 2017-07-04] [WebCite Cache ID 6rhQgTPzi]

9. Scott A, Jeon S, Joyce CM, Humphreys JS, Kalb G, Witt J, et al. A randomised trial and economic evaluation of the effect of response mode on response rate, response bias, and item non-response in a survey of doctors. BMC Med Res Methodol 2011;11:126 [FREE Full text] [doi: 10.1186/1471-2288-11-126] [Medline: 21888678]
10. Zuidgeest M, Hendriks M, Koopman L, Spreeuwenberg P, Rademakers J. A comparison of a postal survey and mixed-mode survey using a questionnaire on patients' experiences with breast care. J Med Internet Res 2011;13(3):e68 [FREE Full text] [doi: 10.2196/jmir.1241] [Medline: 21946048]
11. Statistics Denmark. 2017. The use of ICT in the population in 2015 (IT-anvendelse i befolkningen 2015) URL: https://www.dst.dk/da/Statistik/emner/informationssamfundet/it-anvendelse-i-befolkningen [accessed 2017-05-04] [WebCite Cache ID 6qCnnfxYI]
12. Digitaliseringsstyrelsen [Danish Agency for Digitisation]. Law on Digital Post in Denmark URL: https://www.digst.dk/Servicemenu/English/Policy-and-Strategy/Digital-Post-from-public-authorities [accessed 2017-10-02] [WebCite Cache ID 6tuTzrI7r]
13. borger.dk. 2017. URL: https://www.borger.dk/ [accessed 2017-05-04] [WebCite Cache ID 6qCnvZOF1]
14. e-Boks. 2017. URL: https://www.e-boks.com/danmark/da [accessed 2017-05-04] [WebCite Cache ID 6qCo0za2Q]
15. Digitaliseringsstyrelsen [Danish Agency for Digitisation]. Monthly statistics on digital mail URL: https://www.digst.dk/It-loesninger/Digital-Post/Digital-Post-statistik [accessed 2017-05-04] [WebCite Cache ID 6qCng5IwP]
16. Ritzau F. Newsbreak.dk. 2015 May 16. Danskerne er overraskende gode til at læse digital post fra det offentlige. Folk over 60 år er særligt gode URL: http://newsbreak.dk/overraskende-saa-ofte-bruger-danskerne-e-boks/ [accessed 2017-05-04] [WebCite Cache ID 6qCnRkAja]
17. Moth G, Flarup L, Christensen MB. Kontakt-og sygdomsmønsteret i lægevagten LV-KOS 2011 [Survey on reasons for encounters and disease patterns in OOH primary care LV-KOS 2011]. 2012. URL: http://feap.au.dk/fileadmin/feap/KOS_2008_Rapport_01.pdf [accessed 2018-01-02] [WebCite Cache ID 6wAPC8lch]
18. Callas PW, Solomon LJ, Hughes JR, Livingston AE. The influence of response mode on study results: offering cigarette smokers a choice of postal or online completion of a survey. J Med Internet Res 2010;12(4):e46 [FREE Full text] [doi: 10.2196/jmir.1414] [Medline: 20965873]
19. Zuidgeest M, Hendriks M, Koopman L, Spreeuwenberg P, Rademakers J. A comparison of a postal survey and mixed-mode survey using a questionnaire on patients' experiences with breast care. J Med Internet Res 2011;13(3):e68 [FREE Full text] [doi: 10.2196/jmir.1241] [Medline: 21946048]
20. Horevoorts NJ, Vissers PA, Mols F, Thong MS, van de Poll-Franse LV. Response rates for patient-reported outcomes using web-based versus paper questionnaires: comparison of two invitational methods in older colorectal cancer patients. J Med Internet Res 2015;17(5):e111 [FREE Full text] [doi: 10.2196/jmir.3741] [Medline: 25953059]
21. 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]
22. Sebo P, Maisonneuve H, Cerutti B, Fournier JP, Senn N, Haller DM. Rates, delays, and completeness of general practitioners' responses to a postal versus web-based survey: a randomized trial. J Med Internet Res 2017 Mar 22;19(3):e83 [FREE Full text] [doi: 10.2196/jmir.6308] [Medline: 28330830]
23. Ebert JF, Huibers L, Lippert FK, Christensen B, Christensen MB. Development and evaluation of an “emergency access button” in Danish out-of-hours primary care: a study protocol of a randomized controlled trial. BMC Health Serv Res 2017 May 31;17(1):379 [FREE Full text] [doi: 10.1186/s12913-017-2308-y] [Medline: 28566087]
24. Kantar Media. URL: http://www.kantarmedia.com/global [accessed 2017-06-19] [WebCite Cache ID 6rKaT1z19]

Edited by G Eysenbach; submitted 05.07.17; peer-reviewed by M van Gelder, B Cerutti; comments to author 24.08.17; revised version received 12.10.17; accepted 16.11.17; published 23.01.18.

Please cite as:
Ebert JF, Huibers L, Christensen B, Christensen MB
Paper- or Web-Based Questionnaire Invitations as a Method for Data Collection: Cross-Sectional Comparative Study of Differences in Response Rate, Completeness of Data, and Financial Cost
J Med Internet Res 2018;20(1):e24
URL: http://www.jmir.org/2018/1/e24/
doi:10.2196/jmir.8353
PMID:29362206

Original Paper

Intrapersonal Variation in Goal Setting and Achievement in Health Coaching: Cross-Sectional Retrospective Analysis

Anne M Wallace[1], PhD; Matthew T Bogard[1], MS; Susan M Zbikowski[1], PhD

Wellness Science and Analytics, Humana, Inc, Louisville, KY, United States

Corresponding Author:
Anne M Wallace, PhD
Wellness Science and Analytics
Humana, Inc
500 West Main Street
Louisville, KY, 40202
United States
Phone: 1 513 768 2615
Email: awallace44@gmail.com

Abstract

Background: Chronic conditions in the United States are among the most costly and preventable of all health problems. Research suggests health coaching is an effective strategy for reducing health risks including decreases in weight, blood pressure, lipids, and blood glucose. Much less is known about how and when coaching works.

Objective: The aim of this study was to conduct an analysis of intrapersonal variations in participants' progression in health coaching, examining gender and age-related differences.

Methods: This was a cross-sectional, retrospective analysis of 35,333 health coaching participants between 2012 and 2016. Differences in number of goals and activities set and completed, and number of interactions were assessed using negative binomial models. Differences in goal type were assessed using logistic regression for gender and using the Welch test for age to account for unequal variances.

Results: Participants choosing online coaching were more likely to be younger and female ($P<.001$). Gender and age differences were found for the types of goals set by participants. Regarding program activity, women set and completed 12% more action steps than men ($P<.001$), averaging 21% more interactions than men ($P<.001$); no gender differences were found in number of goals completed ($P=.12$), although the percentage of males and females completing goals was significantly different at 60 and 120 days postenrollment ($P<.001$). Results indicated significant age-related differences in all aspects of program activity: number of interactions, goals set and completed, action steps set and completed (all P values <.01), as well as significant differences in percentage of individuals completing initial goals within 30 days, with older individuals completing more than younger individuals did (all P values <.001).

Conclusions: This study found significant intrapersonal variation in how people participate in and progress through a coaching program. Age-related variations were found in all aspects of coaching activity, from modality preference and initial choice of goal type (eg, weight management, tobacco cessation) to goal completion, whereas gender-related differences were demonstrated for all program activities except number of goals set and completed. These findings indicate that to maximize behavior change, coaches need to personalize the coaching experience to the individual.

(J Med Internet Res 2018;20(1):e32) doi:10.2196/jmir.8892

KEYWORDS

health coaching; health risks; chronic conditions; behavior change

Introduction

Chronic conditions in the United States currently are among the most common, costly, and preventable of all health problems. As of 2012, approximately half of all adults—117 million people—had one or more chronic health conditions [1]. By 2021, according to the Centers for Medicare and Medicaid Services, health care spending will account for almost one-fifth of the gross domestic product [2]. Lifestyle behaviors such as an unhealthy diet, physical inactivity, and tobacco use are among the primary risk factors for disease onset.

Current literature suggests health coaching is an effective strategy for promoting health behavior change [3,4], including improving nutrition, increasing physical activity, and improving adherence to medications [5-7]. Health coaching's effectiveness also has been demonstrated to reduce health risks, including decreases in body mass and weight loss [8-11], positive changes in blood pressure and lipid levels [5,12], and decreases in blood glucose and glycated hemoglobin A_{1c} [13-15]. Moreover, research demonstrates the effectiveness of health and wellness coaching in improving the health status of individuals with chronic conditions, most notably improved self-care regimen compliance [16].

Although the body of literature demonstrating the effectiveness of health coaching is growing, much less is known about how and when it works. Some research has found intrapersonal variation in coaching engagement and retention, including gender and age-related differences [17-19]; however, research has not yet addressed intrapersonal differences in how people set and make progress with goals as part of a coaching program.

Goal setting and achievement are foundational to the coaching process and to health behavior change more generally [20-23]. In particular, behavior change is enhanced via setting of goals that are SMART (specific, measurable, action oriented, realistic, and time bound), accompanied by and supporting short-term goals or action steps [24-29], Moreover, research highlights the importance of obtaining goal commitment as part of the goal-setting process in addition to ongoing monitoring of and review of goals in behavior change interventions [29-31]. For this reason, closer examination of intrapersonal variation in the process of goal activity within the context of health coaching can shed valuable light on how to support behavior change in a variety of different types of people.

In this study, we conduct a detailed analysis of intrapersonal variations in how participants engage in coaching and in their goal-related activities as they progress through a health coaching program, examining gender and age-related differences in the choice of coaching modality, the types of goals set by individuals, and the rate at which goals and supporting action steps are set and completed as participants progress through the program.

Methods

This was a cross-sectional retrospective analysis of individuals enrolled in health coaching as part of an employer-sponsored wellness benefit or as part of wellness programming bundled into an individually purchased health insurance plan. All personally identifiable data were gathered and prepared for analysis following organizational, regulatory, and Institutional Review Board (IRB) policies and practices. The study received IRB approval from Schulman IRB, Cincinnati, OH, on December 12, 2016.

Sample

The sample was comprised of 35,333 individuals aged 18 years or older enrolled in the coaching program between January 2012 and August 2016, and who set one or more goals with their coach. Females comprised the majority of participants, making up 26,778 (75.79%) of the sample. Males comprised 8493 (24.04%) of the sample; the gender of 62 (0.18%) participants was unknown. The age breakdown was as follows: 4653 (13.17%) of participants were younger than 30 years, 18,106 (51.24%) were between 30 and 50 years, 8663 (24.52%) were between 51 and 64 years, and 3911 (11.07%) were 65 years and older.

Intervention

The objective of the coaching program was to reduce health-related risks. Participants could choose to work on one or more health-related areas including weight management, tobacco cessation, healthy eating, fitness, stress management, cholesterol management, diabetes management, blood pressure management, or back care. Participants enrolled in coaching could remain active in the program as long as they were eligible through an employer-sponsored or individual benefit.

Health coaching was delivered via telephone, online, and face-to-face. Face-to-face coaching was available at limited locations and these participants also were able to interact with their coach by telephone and online. All participants enrolled in coaching were given the choice of using either or both telephone and online modalities. Online interactions included both emails from a participant to a coach and journal entries written by a participant to report on progress in coaching; all online correspondence occurred within a HIPPA-secure, password-protected website. Coaches were able to respond to both emails and journal entries.

Goals generally focused on one of the nine health-related areas previously outlined. They were typically set in 30-day increments using SMART format (eg, “I will lose 5 pounds in the next 30 days”). But the goal period/timeframe could be longer or shorter depending on the complexity of the goal and how frequently they interacted with their coach. Once a goal was identified, coaches and participants established action steps to support goal achievement (eg, limiting unhealthy foods to support weight loss, practicing breathing exercises to reduce stress). Supporting activities were most often set in 2 week increments, but could be of shorter duration when appropriate (eg, acquiring exercise equipment or healthy foods). Eligible participants were able to remain in the program as long as they continued to work on setting and achieving goals.

Coaching intervention characteristics were consistent with the components defined by the International Consortium for Health and Wellness Coaching [32], including creating an ongoing relationship with a coach, partnering of coach and individual in setting goals, incorporating self-discovery and active learning processes, and ongoing monitoring of and accountability for progress toward goal completion.

More specifically, the coaching philosophy was holistic and personalized to the individual, designed to facilitate behavior change through a one-to-one relationship with a coach. Coaches had a bachelor's or master's degree in psychology, nutrition, exercise physiology, nursing, or other health profession, and received extensive training in person-centered coaching strategies, cognitive behavioral techniques, positive psychology strategies, and other behavior change methods. Coaches used behavior change techniques to support participants in collaboratively setting goals and action plans, overcoming barriers, enhancing motivation, and assessing/building on progress. Quality of coaching interactions was monitored and evaluated; all coaches underwent 16 hours of training, passed a practicum, and participated in ongoing continuing training and routine quality assessments to assure that coaching protocols were adhered to.

Measures

Data on gender and age, as well as information about health-related status and behaviors (weight, tobacco use, eating habits, stress) were collected during program registration. To better understand age-related trends, age ranges were collapsed into four groups: participants younger than 30 years, those aged between 30 and 50 years, those aged between 51 and 64 years, and those aged 65 years or older.

Coaching Modality

Coaching modality was identified by the types of interactions between coaches and participants documented within the coaching platform. Modality was classified into four groups: (1) online participants who were coached solely via the website, (2) mixed modality participants who worked with their coach by telephone and online, (3) telephone participants who interacted with their coach solely by phone, and (4) face-to-face participants, a combined group who held one or more face-to-face interactions with their coach and may or may not also have worked with their coach online and/or by telephone. Total interactions by method were computed. In all analyses, three online interactions were considered to equal one telephone or face-to-face coaching session; this ratio was derived from subject matter experts independent of the research team.

Goal Type

The coach used a standard list to document a participant’s goal. Categories included weight management, nutrition, fitness, tobacco cessation, stress management, diabetes management, cholesterol management, blood pressure management, back care, or “other.”

Goals Set and Completed

Within the coaching system, coaches documented each goal that was set and completed. The number documented within the system was used for analyses.

Action Steps Set and Completed

Also within the coaching platform, coaches documented each action step set and completed by participants, and the number documented within the system was used for analyses.

Statistical Analysis

Descriptive statistics were generated and significance tests were conducted to test gender and age differences across various measures related to engagement and progress in coaching for participants who enrolled in coaching between January 2012 and August 2016. All analyses were conducted using SAS version 9.4.

Gender and Age Differences in Coaching Modality

Unadjusted multinomial logistic regression was used to model differences in coaching modality (electronic/Web/email, telephone, in-person, mixed) by gender. Linear regression with robust standard errors was used to assess differences in age by modality.

Gender and Age Differences in Goal Type

To assess who set what types of goals, each participant’s goal history was coded to determine if a certain goal type was set (eg, a value of “1” was assigned if a participant set that type of goal at any point in the program; otherwise “0” was assigned). Because participants could set more than one type of goal, separate unadjusted logistic regression models were used to assess the differences in this outcome by gender for each goal type. To assess differences in goal types set by age, mean age was compared for those participants who set a particular goal type (eg, weight management) versus those that never set that type of goal using a Welch test to account for unequal group variances.

Gender and Age Differences in Number of Goals, Action Steps, and Interactions

Intrapersonal variations in the number and type of coaching interactions as well as differences in total goals set and how goals were achieved through action steps were modeled as counts. To determine age and gender differences in the action steps set and completed, as well as goals set and completed, unadjusted negative binomial regression was used. Negative binomial regression models relax the assumption of equidispersion characteristic of a Poisson process.

Gender and Age Differences in Timing of Initial Goal Completion

We determined whether a member completed a goal within 30, 60, or 120 days rather than conduct survival analysis because our data source captured only time to completion for members completing goals. Separate uncontrolled logistic regressions were estimated with completion in 30, 60, or 120 days modeled as binary outcomes.

Results

Modality Preference

Gender differences in the type of coaching interactions chosen were found for all modalities (Table 1). Women were more likely than men were to choose online interactions and engage in face-to-face coaching sessions, whereas men were more likely to choose telephone sessions and engage with their coach via mixed modalities (a combination of telephone and online interactions). Similarly, age-related differences in the type of coaching interactions chosen were found for all comparisons except for online versus in-person interactions (Table 2).

Table 1. Gender differences in modality preference and type of goal set (N=35,271).

Modality and goal type	Male (n=8493)	Female (n=26,778)	Difference[a]	OR (95% CI)[b]	*P*
Modality, predicted probability					
Online	0.63	0.71	–0.08		<.001
Telephone	0.20	0.16	0.05		<.001
In-Person	0.03	0.04	0.02		<.001
Mixed	0.14	0.13	0.01		<.001
Goal type, n (%) [c]					
Weight management	4464 (52.56)	16,596 (61.98)		0.68 (0.65-0.71)	<.001
Fitness	1832 (21.46)	5747 (21.57)		1.01 (0.95-1.07)	.83
Nutrition	1138 (13.40)	3985 (14.88)		0.86 (0.82-0.95)	<.001
Stress management	576 (6.78)	2285 (8.53)		0.78 (0.71-0.86)	<.001
Tobacco cessation	870 (10.24)	1657 (6.19)		1.73 (1.59-1.89)	<.001
Cholesterol	352 (4.14)	738 (2.75)		1.53 (1.34-1.74)	<.001
Blood pressure	407 (4.79)	639 (2.38)		2.06 (1.81-2.34)	<.001
Diabetes	330 (3.88)	678 (2.53)		1.56 (1.36-1.78)	<.001
Back care	206 (2.43)	328 (1.22)		2.01 (1.68-2.39)	<.001

[a] Differences were differences in predicted probabilities from multinomial logistic regression with bootstrapped standard errors.

[b] Odds ratios from unadjusted logistic regression.

[c] The percentage is derived from the total number of goals in each goal type set by each gender.

The average age of members preferring telephone coaching was oldest, approximately 16 years older than the average of those preferring online coaching and 10 years older than those choosing mixed telephone and online interactions.

Type of Goal Set

Differences by gender were found for the types of goals participants chose to set except fitness (Table 1). Of particular note, women were more likely than men were to set goals to manage weight, whereas men were more likely than women were to set goals to quit tobacco or manage their diabetes. Although age differences were found for all goal types (Table 2), the greatest differences were that older individuals were more likely to set condition-related goals (eg, diabetes management, blood pressure management).

Number of Goals, Action Steps, and Interactions

Results indicated no significant gender differences in goal setting; almost half of participants set one goal, with approximately 30% (10,528/35,333, 29.80%) setting two goals and slightly more than 20% (7904/35,333, 22.37%) setting three or more goals (Table 3). Genders differed in other aspects of the coaching process: women set and completed more action steps, and interacted more frequently with their coaches. Women set 12% more action steps than men did, had a mean of 21% more interactions than men did, and completed 12% more action steps than men did. Despite this increased activity among women, no differences by gender were found in overall number of goals goal completed by men and women.

Results indicated significant age-related differences as well (Tables 4 and 5), with older participants generally demonstrating more program activity across all age categories. This trend was apparent in the percentage of participants setting different numbers of goals; among those younger than 30 years, for example, more than 50% (2500/4653, 53.73%) set one goal and less than 15% (695/4653, 14.94%) set three or more goals, whereas among those aged 51 to 64 years, well over half set two or more goals (4842/8663, 55.89%) and 30% (1185/3911, 30.30%) of individuals aged 65 years and older set three or more goals. Significant differences were found among all age-related pairwise comparisons in number of goals set. Older participants set significantly more goals and action steps than their younger counterparts when comparing all age ranges. Coaching interactions peaked among those aged 51 to 64 years, with lower levels among both the older and younger groups. Similar to setting goals and action steps, completion of these coaching activities was highest among the oldest group of participants and decreased significantly at each age range.

Differences in Timing of Initial Goal Completion

Regarding gender, the percentage of males and females completing their first goal within 30 days was approximately 10% (3544/35,271, 10.05%), whereas more than 80% (30,488/35,271, 86.43%) completed their initial goal within 60 days, and almost all participants completed them within 120 days (33,538/35,271, 95.09%) (Table 3). Statistically significant gender differences in the percentage of participants completing their initial goal emerged at 60 and 120 days, Age variations were also seen in the percentage of participants completing their initial goal within 30 days; all pairwise comparisons were significant, with older participants more likely to complete than younger participants were.

This trend continued for initial goal completion at 60 and 120 days, although it was not significant for all pairwise comparisons (Tables 4 and 5).

Goal Completion by Goal Type

To further understand subgroup differences by goal type, the number of goals completed for each of the three most prevalent goal types (weight management, fitness, nutrition)—comprising approximately 80% of all goals set—was also compared. No gender differences were found in number of goals completed within specific goal types (Table 3). Age-related differences were found for all comparisons except those between individuals aged 51 to 64 years and 65 years or older working on fitness or nutrition goals, with individuals of older ages completing more goals within the specific goal types (Tables 4 and 5).

Table 2. Age differences in modality preference and type of goal set (N=35,333).

Modality and area of focus	Age (years)	Comparison[a] age (years)	Difference (SE)	*P*
	Mean (SD)	Mean[b] (SD)		
Modality				
Online	42.27 (11.78)			
Telephone	58.26 (14.74)			
In-person	41.72 (10.65)			
Mixed	49.88 (14.42)			
Modality comparison				
Online vs telephone			15.99 (0.22)	<.001
Online vs in-person			0.55 (0.65)	>.99
Online vs mixed			–7.61 (0.24)	<.001
Telephone vs in-person			16.54 (0.68)	<.001
Telephone vs mixed			8.38 (0.31)	<.001
In-Person vs mixed			–8.16 (0.68)	<.001
Area of focus [c]				
Weight management	45.14 (13.62)	45.70 (14.50)		<.001
Fitness	44.77 (13.87)	45.66 (14.01)		<.001
Nutrition	44.56 (14.42)	45.62 (13.90)		<.001
Stress management	46.30 (14.74)	45.40 (13.91)		.002
Tobacco cessation	44.36 (12.93)	45.56 (14.06)		<.001
Cholesterol	49.24 (13.15)	45.35 (13.99)		<.001
Blood pressure	50.08 (14.39)	45.33 (13.95)		<.001
Diabetes	58.27 (13.07)	45.09 (13.83)		<.001
Back care	52.28 (16.04)	45.37 (13.93)		<.001

[a] Comparisons based on Welch test.

[b] Mean comparison age is the mean age of all participants not working on the designated goal type.

[c] Coaching area of focus includes all participants working on the designated goal type.

Table 3. Gender differences in program activity (N=35,333).

Program activity	Female	Male	OR/exp(β) (95% CI)[a]	*P*
Number of goals set, n (%)				
1 goal	12,701 (47.43)	4175 (49.16)		
2 goals	8062 (30.11)	2456 (28.92)		
≥3 goals	6015 (22.46)	1862 (21.92)		
Activity, mean (SD)				
Number of goals set	2.25 (2.48)	2.22 (2.57)	1.02 (0.98-1.04)	.13
Number of action steps set	5.86 (12.45)	5.25 (11.32)	1.12 (1.07-1.17)	<.001
Number of interactions	5.65 (11.60)	4.69 (10.25)	1.21 (1.16-1.25)	<.001
Number of action steps completed	4.72 (11.94)	4.20 (10.82)	1.12 (1.06-1.19)	<.001
Number of goals completed	1.23 (2.58)	1.19 (2.65)	1.04 (0.99-1.08)	.12
Goals completed, mean (SD)				
Weight management	1.09 (2.25)	1.10 (2.26)	0.99 (0.97-1.02)	.77
Nutrition	1.25 (2.87)	1.20 (2.62)	1.02 (0.97-1.08)	.43
Fitness	1.21 (2.54)	1.16 (2.59)	1.02 (0.96-1.08)	.51
First goal completed, n (%)				
Within 30 days	2644 (9.87)	900 (10.60)	0.92 (0.84-1.02)	.16
Within 60 days	23,027 (85.96)	7461 (87.85)	0.85 (0.78-0.93)	<.001
Within 120 days	25,401 (94.86)	8137 (95.81)	0.72 (0.70-0.93)	.001

[a] Odds ratio (OR) from unadjusted logistic regression for activity and days first goal completed within. Exponentiated coefficients (incident rate ratios) from unadjusted negative binomial regression for number of goals completed.

Table 4. Differences in program activity by age range (N=35,333).

Program activity	Age range (years)			
	<30	30-50	51-64	≥65
Number of goals set, n (%)				
1 goal	2500 (53.73)	8987 (49.64)	3821 (44.11)	1593 (40.73)
2 goals	1458 (31.33)	5382 (29.72)	2555 (29.49)	1133 (29.97)
≥3 goals	695 (14.94)	3737 (20.64)	2287 (26.40)	1185 (30.30)
Activity, mean (SD)				
Number of goals set	1.85 (1.59)	2.12 (2.25)	2.48 (2.89)	2.81 (3.57)
Number of action steps set	3.34 (7.08)	5.03 (11.05)	6.83 (13.82)	9.36 (16.96)
Number of interactions	3.42 (7.99)	5.12 (10.79)	6.79 (13.41)	6.20 (11.37)
Number of action steps completed	2.44 (6.65)	3.96 (10.54)	5.66 (13.29)	7.84 (16.24)
Number of goals completed	0.76 (1.59)	1.09 (2.25)	1.49 (2.89)	1.83 (3.57)
Goals completed, mean (SD)				
Weight management	0.66 (1.45)	0.97 (1.95)	1.30 (2.49)	1.74 (3.41)
Nutrition	0.75 (1.73)	1.16 (2.38)	1.60 (3.76)	1.65 (3.81)
Fitness	0.74 (1.62)	1.05 (2.38)	1.60 (3.16)	1.95 (2.94)
First goal completed, n (%)				
Within 30 days	302 (6.49)	1678 (9.27)	958 (11.06)	621 (15.87)
Within 60 days	4130 (88.76)	15,741 (86.94)	7331 (84.62)	3341 (85.43)
Within 120 days	4462 (95.90)	17,279 (95.43)	8146 (94.03)	3713 (994.94)

Table 5. Comparison of differences in program activity by age range (N=35,333).

Program activity	≥65 vs						51-64 vs				30-50 vs	
	51-64	*P*	30-50	*P*	<30	*P*	30-50	*P*	<30	*P*	<30	*P*
Activity, exp(β) (95% CI) [a]												
Number of goals set	1.14 (1.09-1.18)	<.001	1.33 (1.28-1.37)	<.001	1.52 (1.45-1.60)	<.001	1.17 (1.13-1.20)	<.001	1.34 (1.29-1.40)	<.001	1.15 (1.11-1.19)	<.001
Number of action steps set	1.37 (1.26-1.49)	<.001	1.86 (1.72-2.01)	<.001	2.81 (2.54-3.09)	<.001	1.36 (1.28-1.44)	<.001	2.05 (1.88-2.22)	<.001	1.51 (1.40-1.62)	<.001
Number of interactions	0.91 (0.85-0.98)	.01	1.21 (1.14-1.29)	<.001	1.82 (1.67-1.97)	<.001	1.33 (1.26-1.39)	<.001	1.98 (1.85-2.42)	<.001	1.49 (1.40-1.59)	<.001
Number of action steps completed	1.39 (1.24-1.56)	<.001	1.98 (1.78-2.19)	<.001	3.21 (2.81-3.66)	<.001	1.43 (1.32-1.54)	<.001	2.31 (2.07-2.59)	<.001	1.62 (1.47-1.79)	<.001
Number of goals completed	1.23 (1.13-1.35)	<.001	1.68 (1.55-1.83)	<.001	2.42 (2.18-2.68)	<.001	1.37 (1.29-1.45)	<.001	1.96 (1.80-2.15)	<.001	1.44 (1.32-1.56)	<.001
Goals completed, exp(β) (95% CI) [a]												
Weight management	1.34 (1.22-1.47)	<.001	1.79 (1.64-1.95)	<.001	2.62 (2.35-2.93)	<.001	1.33 (1.42-1.23)	<.001	1.96 (1.78-2.15)	<.001	1.46 (1.34-1.60)	<.001
Nutrition	1.03 (0.85-1.24)	>.99	1.42 (1.20-1.69)	<.001	2.18 (1.78-2.68)	<.001	1.38 (1.23-1.55)	<.001	2.12 (1.8-2.49)	<.001	1.53 (1.33-1.77)	<.001
Fitness	1.22 (0.99-1.51)	.40	1.86 (1.54-2.25)	<.001	2.65 (2.11-3.33)	<.001	1.52 (1.31-1.77)	<.001	2.17 (1.78-2.64)	<.001	1.43 (1.20-1.69)	.001
First goal completed, OR (95% CI) [b]												
Within 30 days	1.52 (1.31-1.76)	<.001	1.85 (1.62-2.11)	<.001	2.72 (2.24-3.30)	<.001	1.22 (1.09-1.36)	<.001	1.79 (1.49-2.15)	<.001	1.47 (1.24-1.75)	.001
Within 60 days	1.07 (0.92-1.23)	>.99	0.88 (0.77-1.01)	.07	0.74 (0.63-0.88)	<.001	0.83 (0.75-0.91)	<.001	0.70 (0.60-0.81)	<.001	0.84 (0.74-0.97)	.005
Within 120 days	1.19 (0.95-1.49)	.25	0.90 (0.72-1.11)	>.99	0.80 (0.61-1.06)	.21	0.75 (0.65-0.88)	<.001	0.67 (0.54-0.85)	<.001	0.89 (0.72-1.11)	>.99

[a] Exponentiated coefficients (incident rate ratios) from unadjusted negative binomial regression. Comparisons were produced using the SAS GENMOD procedure specifying the negative binomial distribution and LSMEANS statement with DIFF, ADJUST, and EXP options.

[b] Odds ratio (OR) from unadjusted logistic regression.

Discussion

In this study of intergroup variations in coaching program participation, we found significant gender- and age-related differences in how people participate in and progress through a coaching program. Age-related variations encompassed all aspects of coaching activity, from initial choice of coaching modality (online, telephone) and goal type (eg, weight management, tobacco cessation) to goal completion as well as time to goal completion, whereas gender-related differences were demonstrated for all program activities except number of goals set and completed.

This research extends previous work indicating intrapersonal variation in program enrollment, retention, and completion. Prior research found gender differences in program engagement and retention in coaching programs [17-19]. Similarly, this study found that women were more likely to engage or interact with their coach than men were. In addition, we found that men and women differ in the modality by which they choose to interact with their coach, with women preferring online interactions, whereas men preferred other forms of interaction. Genders also differed in what they choose to address in coaching; women were more likely to set goals to manage weight, whereas men were more likely to set goals to quit tobacco or manage a condition such as diabetes. Women enrolled in coaching also were more actively involved, not just in interacting more frequently with their coach, but also in setting and completing more action steps. This finding suggests an opportunity to engage men differently in lifestyle change programs. Women and men, however, did not differ in number of goals set and completed.

This work extends past research examining age-related differences among participants in coaching programs, which found variations particularly in program retention and completion [17-19]. In addition to confirming age-related differences in program engagement, this study also found systematic age-related differences in all aspects of program activity. Participants preferring telephone coaching were, on average, 15 years older than those preferring to interact with their coach online. Age-related variations also were found among the types of goals participants chose to set, most notably with older participants being more likely to work on goals that support management or reduction of health-related risks such as elevated blood pressure, cholesterol, or diabetes. Additionally, all aspects of program progression—setting goals and action steps as well as completing them—saw increasing rates of activity with increasing age. In particular, goal completion increased with age across all goal types as well as within specific areas of focus (eg, weight management, nutrition, fitness). Similar age-related trends were seen in the percentage of participants completing their initial goal, with significant variation in completion at 30 days and some variation depending on age comparisons at 60 and 120 days postenrollment. Only among program interactions did the trend of increased activity with age vary somewhat, with number of coach-participant interactions peaking among those in the 51 to 64 years age group and declining somewhat among those 65 years and older.

These findings may offer new insights to help better design and target wellness promotion and interventions that lead to behavior change and health improvement. Results of this study underscore the importance of addressing intrapersonal differences. Starting with promotional materials, individuals of different ages and genders may respond more favorably to messaging tailored to their preferred areas of focus (eg, weight loss, tobacco cessation); alternatively, organizations could shift their messaging to entice enrollment in coaching for areas not currently utilized as heavily. Once enrolled in coaching, coaches may need to work more actively to engage men and younger participants in various aspects of the coaching process, providing additional support around setting and completing action steps to support goals with the knowledge that completing more action steps increases the likelihood of goal completion.

Finally, despite intrapersonal variation, coached participants continue to have much in common. For example, the majority of participants in coaching chose to work on weight management despite significant differences in other areas of focus. Likewise, increased rates of action step completion promote goal completion, regardless of gender or age. These findings strongly indicate that the process coaches use when working with participants should remain structured yet flexible, providing a framework setting the stage for behavior change while also personalizing the experience on the individual to meet his or her unique needs.

Strengths and Limitations

This has many strengths, which include evaluating a large national sample with demographic and operational data from a diverse set of employers offering the same health coaching program to their employees. With these strengths, there are some limitations to point out.

First, results may only generalize to employer-sponsored health coaching programs and not to other types of wellness programs (non-employer sponsored program) or to other populations such as Medicare, etc. Additionally, this study included two key demographic metrics, age and gender, but did not include race or socioeconomic indicators because these were not collected. Information regarding chronic conditions was also not available for this study. Additional patterns and findings could be uncovered with additional demographic and condition-related data.

Health coaching programs offered may differ in the modalities delivered, length of treatment, etc. Thus, results may not generalize to other health coaching programs offered to employers. However, this program included the core elements defined by the International Consortium for Health and Wellness Coaching and should generalize to others meeting these standards.

Future Directions

Expanding this work in several ways can widen its applicability within the coaching process. In this study, we explored how intrapersonal demographic factors influence variations in coaching participation and progress. Additional work is needed around psychological and behavioral factors and how they influence coaching participation and progress, as well as environmental and cultural factors within the worksite and beyond. Our findings, for example, suggest that if we can find new and different ways to engage younger participants, who may not yet feel the need for lifestyle change, we may inculcate healthy behaviors at a younger age and potentially reduce the need for people to address chronic health-related risks later in life. Alternatively, younger individuals may be more amenable to primarily digital programs and/or programming that incorporates social media. Supplemental work identifying these and other factors can provide a more holistic picture of the influencers of participation and progress in wellness programming.

Additionally, it will be important to connect this work to program outcomes beyond goal completion or program completion. Examination of health-related outcomes, such as weight loss and positive biometric changes, as well as the subjective appraisal of health are important to understanding the influence of intrapersonal variations on health status in addition to their influence of program participation and progression.

Conclusions

Research in health coaching demonstrates it is a key intervention in health behavior change, and that the process of goal setting and achievement is foundational to the intervention's success. The question of how to optimize coaching interventions, however, requires significant additional study. This study found significant intrapersonal variation in how people participate in and progress through a coaching program. Age-related variations were found in all aspects of coaching activity, from modality preference and initial choice of goal type (eg, weight management, tobacco cessation) to goal completion, whereas gender-related differences were demonstrated for all program activities except number of goals set and completed. These findings indicate that to maximize behavior change, coaches need to personalize the coaching experience to the individual.

Acknowledgments

We would like to thank research consultant, Varun Kukreja, MBA, for the exploratory analyses that led to this study; Laura Happe, PharmD, MPH, for reviewing earlier versions of this paper and providing editorial feedback; and Mona Deprey for her assistance reviewing and finalizing the manuscript for publication.

Conflicts of Interest

The authors are employees of Humana, Inc. AMW and SMZ own stock in Humana, Inc. No other conflicts of interest are reported by the authors of this paper.

References

1. Ward B, Schiller J, Goodman R. Multiple chronic conditions among US adults: a 2012 update. Prev Chronic Dis 2014 Apr 17;11:E62 [FREE Full text] [doi: 10.5888/pcd11.130389] [Medline: 24742395]
2. Keehan SP, Cuckler GA, Sisko AM, Madison AJ, Smith SD, Lizonitz JM, et al. National health expenditure projections: modest annual growth until coverage expands and economic growth accelerates. Health Aff (Millwood) 2012 Jul;31(7):1600-1612 [FREE Full text] [doi: 10.1377/hlthaff.2012.0404] [Medline: 22692089]
3. Hill B, Richardson B, Skouteris H. Do we know how to design effective health coaching interventions: a systematic review of the state of the literature. Am J Health Promot 2015;29(5):e158-e168. [doi: 10.4278/ajhp.130510-LIT-238] [Medline: 24720388]
4. Terry PE, Seaverson EL, Grossmeier J, Anderson DR. Effectiveness of a worksite telephone-based weight management program. Am J Health Promot 2011;25(3):186-189. [doi: 10.4278/ajhp.081112-QUAN-281] [Medline: 21192748]
5. Olsen JM, Nesbitt BJ. Health coaching to improve healthy lifestyle behaviors: an integrative review. Am J Health Promot 2010;25(1):e1-e12. [doi: 10.4278/ajhp.090313-LIT-101] [Medline: 20809820]
6. Alley S, Jennings C, Plotnikoff RC, Vandelanotte C. Web-based video-coaching to assist an automated computer-tailored physical activity intervention for inactive adults: a randomized controlled trial. J Med Internet Res 2016 Aug 12;18(8):e223 [FREE Full text] [doi: 10.2196/jmir.5664] [Medline: 27520283]
7. Dennison L, Morrison L, Lloyd S, Phillips D, Stuart B, Williams S, et al. Does brief telephone support improve engagement with a web-based weight management intervention? Randomized controlled trial. J Med Internet Res 2014;16(3):e95 [FREE Full text] [doi: 10.2196/jmir.3199] [Medline: 24681761]
8. Mao AY, Chen C, Magana C, Caballero BK, Olayiwola JN. A mobile phone-based health coaching intervention for weight loss and blood pressure reduction in a national payer population: a retrospective study. JMIR Mhealth Uhealth 2017 Jun 08;5(6):e80 [FREE Full text] [doi: 10.2196/mhealth.7591] [Medline: 28596147]
9. Allman-Farinelli M, Partridge SR, McGeechan K, Balestracci K, Hebden L, Wong A, et al. A mobile health lifestyle program for prevention of weight gain in young adults (TXT2BFiT): nine-month outcomes of a randomized controlled trial. JMIR Mhealth Uhealth 2016 Jun 22;4(2):e78 [FREE Full text] [doi: 10.2196/mhealth.5768] [Medline: 27335237]
10. Bennett GG, Herring SJ, Puleo E, Stein EK, Emmons KM, Gillman MW. Web-based weight loss in primary care: a randomized controlled trial. Obesity (Silver Spring) 2010 Feb;18(2):308-313 [FREE Full text] [doi: 10.1038/oby.2009.242] [Medline: 19696764]
11. Kivelä K, Elo S, Kyngäs H, Kääriäinen M. The effects of health coaching on adult patients with chronic diseases: a systematic review. Patient Educ Couns 2014 Nov;97(2):147-157. [doi: 10.1016/j.pec.2014.07.026] [Medline: 25127667]
12. Newnham-Kanas C, Gorczynski P, Morrow D, Irwin J. Annotated bibliography of life coaching and health research. Int J Evid Based Coach Mentor 2009 Feb;7(1):39-103 [FREE Full text]
13. Quinn CC, Shardell MD, Terrin ML, Barr EA, Ballew SH, Gruber-Baldini AL. Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control. Diabetes Care 2011 Sep;34(9):1934-1942 [FREE Full text] [doi: 10.2337/dc11-0366] [Medline: 21788632]
14. Wolever RQ, Dreusicke M, Fikkan J, Hawkins TV, Yeung S, Wakefield J, et al. Integrative health coaching for patients with type 2 diabetes: a randomized clinical trial. Diabetes Educ 2010;36(4):629-639. [doi: 10.1177/0145721710371523] [Medline: 20534872]
15. Wayne N, Perez DF, Kaplan DM, Ritvo P. Health coaching reduces HbA1c in type 2 diabetic patients from a lower-socioeconomic status community: a randomized controlled trial. J Med Internet Res 2015;17(10):e224 [FREE Full text] [doi: 10.2196/jmir.4871] [Medline: 26441467]

16. Hutchison AJ, Breckon JD. A review of telephone coaching services for people with long-term conditions. J Telemed Telecare 2011;17(8):451-458. [doi: 10.1258/jtt.2011.110513] [Medline: 22025743]
17. Brill PA, Kohl HW, Rogers T, Collingwood TR, Sterling CL, Blair SN. The relationship between sociodemographic characteristics and recruitment, retention, and health improvements in a worksite health promotion program. Am J Health Promot 1991;5(3):215-221. [Medline: 10148671]
18. Grossmeier J. The influence of worksite and employee variables on employee engagement in telephonic health coaching programs: a retrospective multivariate analysis. Am J Health Promot 2013;27(3):e69-e80. [doi: 10.4278/ajhp.100615-QUAN-190] [Medline: 23286600]
19. Terry PE, Fowles JB, Harvey L. Employee engagement factors that affect enrollment compared with retention in two coaching programs--the ACTIVATE study. Popul Health Manag 2010 Jun;13(3):115-122. [doi: 10.1089/pop.2009.0040] [Medline: 20521904]
20. 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]
21. Pearson ES. Goal setting as a health behavior change strategy in overweight and obese adults: a systematic literature review examining intervention components. Patient Educ Couns 2012 Apr;87(1):32-42. [doi: 10.1016/j.pec.2011.07.018] [Medline: 21852063]
22. Ryan P. Integrated Theory of Health Behavior Change: background and intervention development. Clin Nurse Spec 2009;23(3):161-170; quiz 171 [FREE Full text] [doi: 10.1097/NUR.0b013e3181a42373] [Medline: 19395894]
23. Bandura A. The primacy of self-regulation in health promotion. Applied Psych 2005 Apr;54(2):245-254 [FREE Full text] [doi: 10.1111/j.1464-0597.2005.00208.x]
24. Lawn S, Schoo A. Supporting self-management of chronic health conditions: common approaches. Patient Educ Couns 2010 Aug;80(2):205-211. [doi: 10.1016/j.pec.2009.10.006] [Medline: 19931372]
25. Kolehmainen N, Francis JJ. Specifying content and mechanisms of change in interventions to change professionals' practice: an illustration from the Good Goals study in occupational therapy. Implement Sci 2012 Oct 18;7:100 [FREE Full text] [doi: 10.1186/1748-5908-7-100] [Medline: 23078918]
26. Bandura A. Social cognitive theory of self-regulation. Organ Behav Hum Dec 1991 Dec;50(2):248-287 [FREE Full text] [doi: 10.1016/0749-5978(91)90022-L]
27. Elliot DL, Goldberg L, MacKinnon DP, Ranby KW, Kuehl KS, Moe EL. Empiric validation of a process for behavior change. Transl Behav Med 2016 Dec;6(3):449-456 [FREE Full text] [doi: 10.1007/s13142-015-0343-y] [Medline: 27528533]
28. Lentferink AJ, Oldenhuis HK, de Groot M, Polstra L, Velthuijsen H, van Gemert-Pijnen JE. Key components in eHealth interventions combining self-tracking and persuasive eCoaching to promote a healthier lifestyle: a scoping review. J Med Internet Res 2017 Aug 01;19(8):e277 [FREE Full text] [doi: 10.2196/jmir.7288] [Medline: 28765103]
29. Wolever RQ, Caldwell KL, Wakefield JP, Little KJ, Gresko J, Shaw A, et al. Integrative health coaching: an organizational case study. Explore (NY) 2011;7(1):30-36. [doi: 10.1016/j.explore.2010.10.003] [Medline: 21194670]
30. 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]
31. 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]
32. Wolever RQ, Simmons LA, Sforzo GA, Dill D, Kaye M, Bechard EM, et al. A systematic review of the literature on health and wellness coaching: defining a key behavioral intervention in healthcare. Glob Adv Health Med 2013 Jul;2(4):38-57 [FREE Full text] [doi: 10.7453/gahmj.2013.042] [Medline: 24416684]

Abbreviations

IRB: Institutional Review Board
OR: odds ratio

Edited by G Eysenbach; submitted 01.09.17; peer-reviewed by T Bose, M Vollenbroek-Hutten, M Thirumalai; comments to author 28.09.17; revised version received 21.11.17; accepted 08.12.17; published 26.01.18.

Please cite as:
Wallace AM, Bogard MT, Zbikowski SM
Intrapersonal Variation in Goal Setting and Achievement in Health Coaching: Cross-Sectional Retrospective Analysis
J Med Internet Res 2018;20(1):e32
URL: http://www.jmir.org/2018/1/e32/
doi:10.2196/jmir.8892
PMID:29374005

Original Paper

Identifying Opinion Leaders to Promote Organ Donation on Social Media: Network Study

Jingyuan Shi[1], PhD; Charles T Salmon[2], PhD

[1]Department of Communication Studies, Hong Kong Baptist University, Hong Kong, China (Hong Kong)
[2]Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore, Singapore

Corresponding Author:
Jingyuan Shi, PhD
Department of Communication Studies
Hong Kong Baptist University
5 Hereford Road, Communication and Visual Arts Building
Kowloon Tong
Hong Kong,
China (Hong Kong)
Phone: 852 34118137
Email: jolieshi@hkbu.edu.hk

Abstract

Background: In the recent years, social networking sites (SNSs, also called social media) have been adopted in organ donation campaigns, and recruiting opinion leaders for such campaigns has been found effective in promoting behavioral changes.

Objective: The aim of this paper was to focus on the dissemination of organ donation tweets on Weibo, the Chinese equivalent of Twitter, and to examine the opinion leadership in the retweet network of popular organ donation messages using social network analysis. It also aimed to investigate how personal and social attributes contribute to a user's opinion leadership on the topic of organ donation.

Methods: All messages about organ donation posted on Weibo from January 1, 2015 to December 31, 2015 were extracted using Python Web crawler. A retweet network with 505,047 nodes and 545,312 edges of the popular messages (n=206) was constructed and analyzed. The local and global opinion leaderships were measured using network metrics, and the roles of personal attributes, professional knowledge, and social positions in obtaining the opinion leadership were examined using general linear model.

Results: The findings revealed that personal attributes, professional knowledge, and social positions predicted individual's local opinion leadership in the retweet network of popular organ donation messages. Alternatively, personal attributes and social positions, but not professional knowledge, were significantly associated with global opinion leadership.

Conclusions: The findings of this study indicate that health campaign designers may recruit peer leaders in SNS organ donation promotions to facilitate information sharing among the target audience. Users who are unverified, active, well connected, and experienced with information and communications technology (ICT) will accelerate the sharing of organ donation messages in the global environment. Medical professionals such as organ transplant surgeons who can wield a great amount of influence on their direct connections could also effectively participate in promoting organ donation on social media.

(J Med Internet Res 2018;20(1):e7) doi:10.2196/jmir.7643

KEYWORDS

social media; health promotion; organ donation; opinion leaders; social network analysis

Introduction

Organ Donation and Social Media

Since the 1960s, many countries have initiated organ donation programs, and at least eighty countries are now known to have a national organ donation program [1]. However, the organ donation program is still in its infancy in China. In China, the first nationwide organ donation program was not launched until 2013 [2]. This new opt-in organ donation program, however, has met with little success: approximately 1.5 million patients in China need an organ transplant each year, but only 10,000 of them are able to receive one [3]. By August 2017, only

310,620 Chinese people had registered as organ donors [4]. The great shortage of available organs in China underscores the need for organ donation promotion.

Social networking sites (SNSs, also called social media) are a popular platform for promoting organ donation in the United States [5,6] and in some other Western countries [7]. Organ donation campaigns using social media have yielded some promising results such as increasing organ donor consent rates. Weibo is one of the most popular social media platforms in China, with 600 million registered users [8] and is often regarded as the equivalent of Twitter in China. A previous content analysis revealed that people have shared their cultural beliefs, complex feelings, and concerns about organ donation, as well as their willingness to donate on Weibo [9]. Thus, Weibo is a potential platform for promoting organ donation in China. This paper focuses on the dissemination of organ donation tweets on Weibo and examines the opinion leadership in the retweet network of popular organ donation messages on Weibo using social network analysis.

To provide the background for the study, this paper will first review the literature on health communication campaigns and opinion leadership on social media. Following that literature review, this paper's hypotheses are set forth. Finally, the Methods, Results, and Discussion are presented.

Identifying Opinion Leadership for Social Media Health Campaigns

Recruiting opinion leaders has proven to be an effective strategy in Web-based organ donation promotions. For example, Stefanone et al [5] recruited peer leaders to endorse organ donation on social media and reported an increase in organ donor card requests and donor registration among college students at the end of the campaign. Such empirical evidence indicates that opinion leader interventions could encourage organ-donor registration behavior in SNS communities. In fact, opinion leaders have long been regarded as change agents in health campaigns because they are able to assist in the implementation of behavior change efforts: they can legitimize the behavior change program, convey the health messages, and act as role models for behavior change [10]. However, little research to date has identified opinion leaders for organ donation in the context of social media.

The classic two-step flow of communication hypothesis suggests that opinion leaders are individuals who directly receive information from mass media and, in turn, pass on what they know to their everyday associates through interpersonal communication [11,12]. Although this definition concerns opinion leaders' access to information via mass media, subsequent definitions of opinion leaders have focused on the extent of influence, such as the impact on the opinions, attitudes, and behaviors that they exert on others [10,13,14]. Furthermore, recent research has discussed how the notion of opinion leadership on social media has evolved from earlier eras. Nowadays, many individual users—rather than merely a few opinion leaders— have substantial access to information, as well as the ability to share it with mass audiences in an instant [15]. It seems that opinion leaders' privilege in accessing and disseminating information, which is suggested by the "two-step flow" hypothesis, no longer exists in the context of social media. Thus, opinion leadership on social media, may rest not only on the ability to access or disseminate information but also on the ability to bridge groups [14], build critical links in information dissemination [14], and trigger others in the network to share information [16].

Moreover, besides the "two step flow" process [12], other models of information dissemination, including direct one-step and complex network flows, also exist on social media [17,18]. A recent analysis of communication flows on Twitter reveals that general Twitter accounts receive information directly from traditional media and official accounts, as well as indirectly from intermediating amplifiers, who are individual or organizational users with public authority or public visibility [17]. The information dissemination on Twitter also follows a network step-flow that includes the coexistence of one-step flow, two-step flow, and a multi-step back-and-forth flow of communication among media and official accounts, general Twitter accounts, and amplifiers [17]. Thus, under such a complex model of information dissemination on social media, influential users emerge in local contexts, as well as in the overall communication network. In fact, such unique features of opinion leadership on social media are reflected in Bodendorf and Kaiser's [19] recent conceptualization of online opinion leadership, which consists of two dimensions: local and global. Although both dimensions involve the ability of influencing and controlling information flow, they are slightly different from each other. Local opinion leadership refers to influence in a direct but limited environment; for example, a direct influence on one's neighbors. Alternatively, global opinion leadership refers to indirect influence on others during the information exchange. For instance, this could be an ability to control the overall information flow in a whole network. In general, such local and global opinion leaders are crucial for the implementation of successful health promotions and interventions using social media platforms [20].

In terms of the operationalization of opinion leadership, scholars have employed various methods to identify opinion leaders who are able to assist in the implementation of behavior modification efforts [10]. Among all these means, the sociometric method has been regarded as "the most valid and reliable" method [10] and is "more precise than self-designating method" [21]. The sociometric method is able to capture not only direct flows of information but also a completed network of information dissemination and exchange. Within this network, several network metrics can be used to calculate the structural position a member has secured. Of all the network metrics, scholars have used indegree centrality most frequently to measure opinion leadership when employing the sociometric technique [20]. Previous sociometric studies have documented the positive relationships between indegree centrality and both self-reported and other-identified opinion leadership in offline and online environments [22-24]. In a retweet network on social media, indegree represents the number of direct ties a member receives from its neighbors, and members with a large number of indegree are prominent [25]. Hence, indegree centrality is a good indicator of one's local opinion leadership [19]. However, this metric does not measure a member's indirect connections in the network. To capture such indirect influence (ie, global opinion leadership), Bodendorf and Kaiser [19] proposed two other network metrics: closeness centrality and betweenness centrality.

Closeness is the sum of shortest distances from a member to all other members in the network [26], and the normalized version of closeness is divided into n – 1 (n is the number of all members in the network) [27]. In a directed network, this closeness can be calculated for sending (ie, out-closeness) and receiving (ie, in-closeness) [25]. In a retweet network of information on social media, a member with a high in-closeness score secures a position with a short distance to most others. It means that the tweet posted by this member will spread quickly to a random member in the network through network ties. Thus, we employ in-closeness instead of out-closeness as the indicator of opinion leadership for the retweet network in this study. Alternatively, betweenness measures how often a member falls along the shortest path between two other members in the network [26]. In a retweet network, members with high betweenness are usually acting as a bridge that connects different cluster of Weibo users. Members with high betweenness thus play gatekeeper roles in networks that control information flows and facilitate information dissemination beyond the boundaries of local groups [28,29]. Thus, this study adopts the Bodendorf and Kaiser [19] measures and employs three sociometric indicators for opinion leadership in a retweet network: indegree for local opinion leadership, and in-closeness and betweenness for global opinion leadership.

The Predictors of Opinion Leadership

In additional to the indicators of opinion leadership, scholars have also long investigated the factors associated with opinion leadership. In 1957, Katz's [11] classical article on opinion leaders proposed three predictors of personal influence: (1) personal attributes (who one is), (2) competence (what one knows), and (3) social position (whom one knows). This paper examines how these three factors are associated with opinion leadership for organ donation on a Chinese SNS. In this research, personal attributes include one's activeness, verification status (ie, if the user is a verified account on Weibo), and geographic location; competence refers to one's knowledge about medical issues in general; and social position refers to the numbers of followers and followings the user has on the SNS.

Personal Attributes

Activeness

Sociability is the first factor related to leadership in offline and online contexts [30]. Individuals who engage in more communication activities can more easily obtain information and build relationships [31], thus having more potential to extend their reach to and influence others [32]. In the SNS context, research has found that one's activeness on social media was positively associated with the number of retweets his or her posts received [33] and the probability of building communication ties with others [34]. Hence, the following hypothesis is proposed:

> Hypothesis 1 (H1): a user's activeness on an SNS is positively associated with his or her (1) local and (2) global opinion leadership for organ donation on social media.

Verification

On social media, some users' accounts include a verified badge on their profile that shows the authenticity of their identities as key individuals or organizations. To obtain the verified badge, a verification request is usually submitted by the user and then confirmed by the SNS platform. Previous research has stated that the verification badge indicates a user's credibility [33] or eliteness [35] on the SNS. For example, Zhang et al [33] claimed that verified accounts were perceived to be more credible than unverified accounts and found that messages posted via verified accounts attracted a larger number of retweets than did the messages posted via unverified accounts. Thus, having a verified badge on one's profile on social media may lead to more influence on the SNS platform:

> Hypothesis 2 (H2): compared with unverified users, verified users exhibit more (1) local or (2) global opinion leadership for organ donation on social media.

Information and Communication Technology Development in One's Location

In addition to being associated with one's activeness and identification on the SNS platform, opinion leadership in cyberspace may be subject to certain external environmental factors such as opportunities to access information and communications technologies (ICTs), including the Internet, cell phones, and personal digital assistants [36]. Opportunities for accessing and utilizing ICTs in one's geographic location are closely tied to an individual's ability to develop and maintain social relationships online [37,38]. Indeed, Lyons and Henderson [39] revealed that computer skill and Internet self-efficacy are positively associated with a person's opinion leadership in a computer-mediated environment. Moreover, a recent study on users' influence in SNS communities found that, within a Weibo community about human immunodeficiency virus or acquired immune deficiency syndrome, users coming from areas with well-developed ICTs secured more influential positions in their follower-following network than did those who came from areas with underdeveloped ICTs [36]. Thus, it is reasonable to expect a positive effect of the level of ICT development in one's location on his or her influence on the organ donation topic in virtual communities:

> Hypothesis 3 (H3): the degree of ICT development in a user's location is positively associated with his or her (1) local and (2) global opinion leadership for organ donation on social media.

Competence

An individual's expertise or knowledge about a social issue has long been regarded as a critical contributor to his or her stature as an opinion leader on the topic [11,32]. The research on product diffusion has found that opinion leaders are more knowledgeable about the product than nonleaders and that individuals who are superior in professional knowledge are also more likely to become opinion leaders in computer-mediated environments as well [39]. In terms of organ donation, medical knowledge is critical to improving the willingness of donation and reducing refusal from potential donors' relatives [40]. On Weibo, physicians are a special group of users. Some of them not only received verification from the platform but also had detailed profile information that showed their professional position in a clinic, hospital, or university. Thus, this group of users who possess medical knowledge are considered credible and authoritative and exhibit more opinion leadership for organ donation:

> Hypothesis 4 (H4): users with a medical-focused profile on social media exhibit more (1) local and (2) global opinion leadership for organ donation than others without medical-focused profiles.

Social Positions

SNS users not only integrate their offline social relationships into a cyber network but also develop new online social ties on social media [15]. In a network of social relationships on social media, a user disseminates information directly to his or her followers. Meanwhile, the user himself or herself is exposed to the information his or her connections send out. Having a large number of followers (those who followed the focal user) enables a user to disseminate information efficiently, whereas a large number of followings (those whom followed by the focal user) provides the user with a broad source of information. Indeed, previous research has shown that individuals who are well connected on social media are more influential than others in the virtual environment [18]. For example, Zhang et al [33] found that the number of followers a person had was positively associated with the number of retweets and comments that person's posts received. Thus, users with numerous followers and followings obtain a well-connected location and are influential in the SNS network:

> Hypothesis 5 (H5): the number of followers a user has is positively associated with his or her (1) local and (2) global opinion leadership for organ donation topic on social media.
>
> Hypothesis 6 (H6): the number of followings a user has is positively associated with his or her (1) local and (2) global opinion leadership for organ donation on social media.

Methods

Data Collection and Sample

Weibo is a Twitter-like microblogging service site that was launched in 2009 and that has become one of the most popular social media platform in China, with 600 million registered users [8]. This study used the built-in Weibo main search function to retrieve all messages about organ donation posted from January 1, 2015 to December 31, 2015. The search keywords included "donation/donating/donated" and the names of organs and tissues listed on the Chinese organ donor registration application (including organ/organs, body/bodies, kidney/kidneys, liver/livers, heart/hearts, lung/lungs, pancreas, small intestines, and cornea). Python Web crawler was employed to extract all available searching results on January 10, 2016, which consisted of 7465 Weibo posts. To eliminate irrelevant messages, the Weibo posts were manually coded into two categories: (1) relevant, which discusses organ donors, organ recipients, organ donation systems, or organ donation policy and ethics and (2) irrelevant, which discusses issues not relevant to the aforementioned organ donation topics. Two native Chinese speaker coders independently coded 10.04% (750/7465) of the posts, randomly chosen from the data, and the intercoder reliability Cohen kappa [41] was .92. Next, all the 7465 posts were split in half and separately coded by the two coders, and 6701 messages (89.77%, 6701/7465) were coded as organ donation messages.

The popular organ donation messages were defined as the messages whose number of retweets ranked in the top three percentiles out of all 6701 messages (n=206). The retweet network of the 206 popular messages was extracted using Python Web crawler in April 2016, resulting in a retweet network with 505,047 unique Weibo users. Next, Python Web crawler was employed to extract the profile information of the Weibo users who received at least one retweet from others in the retweet network (n=44,074). The Python Web crawler extracted all existing accounts as of April 2016, which included 43,510 users. The information about these users' profiles included the account's username, verification status, self-introduction, self-reported location and gender, as well as his or her number of followers, followings, and posts on Weibo.

Constructing the Retweet Network

The retweet network of the popular organ donation messages was constructed such that if user *i* retweets the post of member *j*, then *i* was connected to *j*. The direction of the tie was from *i* to *j*, with a weight that equals the number of times that *i* retweets posts from *j*. Thus, the retweet network is a directed, weighted network. We constructed and analyzed the retweet networks using the igraph package [42] in R.

Measures and Analytical Design

Opinion Leadership

An individual's opinion leadership was measured via three network metrics from the retweet network, including his or her indegree for local opinion leadership, as well as in-closeness and betweenness for global opinion leadership. The indegree ranged from 1 to 59,061 with a mean of 12.51 (standard deviation [SD] 515.32). The in-closeness ranged from 3.92e-12 to 6.20e-12 with a mean of 3.95e-12 (SD 1.17e-13). The betweenness ranged from 0 to 58,442,335.20 with a mean of 80,485.50 (SD 939,107.11).

Activeness

The measurement of a user's activeness on Weibo was adapted from Zhang et al [33], which used the total number of messages a user has posted. Among the 43,510 users, the number of posts ranged from 0 to 677,495 with a mean of 11,170.01 (SD 16,617.82).

Verification

If a user had a verified badge in his or her Weibo profile, this account was regarded as a verified account. It was a dichotomous variable: *verified*=1 (4158/43,510, 9.5%) and *unverified*=0 (39,352/43,510, 90.4%).

Table 1. The correlation matrix among all continuous variables.

Variable	1	2	3	4	5	6	7
1. Indegree	–						
2. In-closeness	0.266[a]	–					
3. Betweenness	0.117[a]	0.567[a]	–				
4. The number of followers	0.587[a]	0.163[a]	0.043[a]	–			
5. The number of followings	0.068[a]	0.115[a]	0.076[a]	0.133[a]	–		
6. Activeness on Weibo	0.117[a]	0.163[a]	0.200[a]	0.171[a]	0.361[a]	–	
7. Location	0.037[a]	0.043[a]	0.041[a]	0.045[a]	0.027[a]	0.111[a]	–

[a] *P*<.01.

Information and Communication Technology Development

The user’s location was recoded as a continuous variable according to its degree of ICT development. According to China’s ICT development index [43] and the digital access index for all countries in the world [44], the location was coded into a continuous variable from 1 to 6, with a higher value representing a better degree of ICT development (mean=3.83, SD=1.29).

Medical-Focused

If a user has a medical-focused profile which lists the user’s professional position in a clinic, hospital, or university, this variable was coded as 1 (n=84).

The Numbers of Followers and Followings

This information was listed in users’ profiles. The number of followers of all the users ranged from 0 to 50,563,948 with a mean of 38,801.45 (SD 827,423.30). The number of followings ranged from 1 to 5832 with a mean of 543.84 (SD 545.83).

Analytical Design

The general linear model (GLM) was employed for analysis. Due to highly skewed distributions, indegree, in-closeness, betweenness, activeness, and the numbers of followers and followings were square root transformed to meet the assumptions of GLM. The correlation matrix among continuous variables is presented in Table 1.

Results

Mapping the Retweet Network of Popular Organ Donation Messages

The retweet network is a connected network including 505,047 nodes and 545,312 edges. The length of the maximum distance between nodes (ie, diameter) in this network is 21. The network has a low density: only 0.0002% of possible edges between all the nodes are connected. The reciprocity values and clustering coefficient of this network are extremely low at 0.0016 and 0.000003, respectively. The low values of density, reciprocity, and clustering coefficient indicate a sparse network. In addition, indegree centralization is 0.1169, and outdegree centralization is 0.00017 for this network. This indicates that, in terms of indegree, links are retweeted disproportionately to a small group of users. The distribution of indegree within this retweet network is highly skewed (see Figure 1). All the network-level statistics indicate that this retweeting network is a sparse and centralized network. Table 2 summarizes the network-level statistics. Figure 2 visualizes this retweet network with nodes whose indegree is equal or larger than 50.

Table 2. The network-level characteristics of the retweeting network.

Social network metric	Definition	Possible range	Value
Size	The number of nodes (eg, users) in the network	N/A[a]	505,047
Diameter	The largest geodesic distance, which is the shortest distance from one node to another in the network	N/A	21
Density	The proportion of all possible dyadic connections that are presented in the network	0-1	0.000002
Reciprocity	The proportion of all pairs in the network that have a reciprocated tie between them	0-1	0.0016
Clustering coefficient	The degree to which nodes in the network tend to cluster together	0-1	0.000003
Indegree centralization	The extent to which the distribution of indegree centrality in the network deviates from a perfectly equal distribution	0-1	0.1169
Outdegree centralization	The extent to what the distribution of outdegree centrality in the network deviates from a perfectly equal distribution	0-1	0.00017

[a] N/A: not applicable.

Figure 1. The log transformed distribution of indegree.

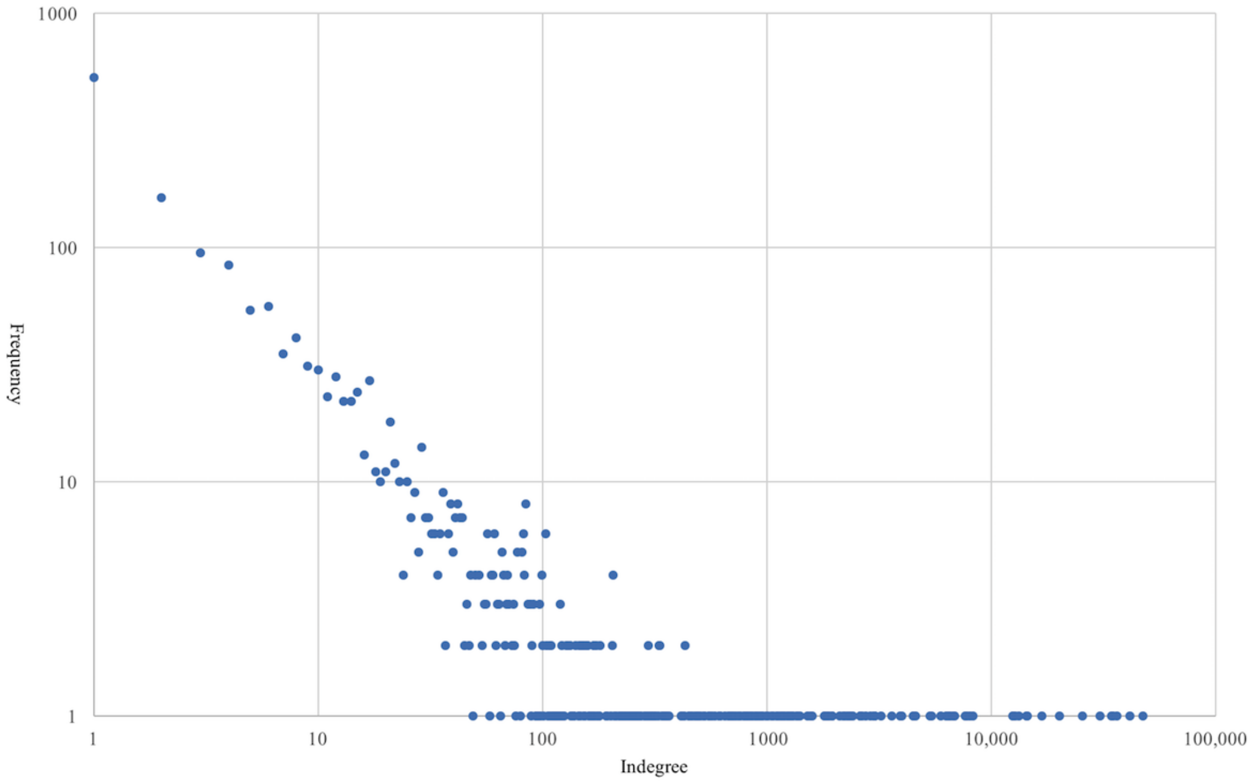


Figure 2. Visualization of the sharing network of popular organ donation messages. The figure includes only nodes whose indegree equal or larger than 50. Nodes represent Weibo users (n=362). The size of node depends on its indegree. The larger the node, the greater amount of retweets the user received. Lines represent retweet relationship between Weibo users. The weight of line depends on the number of retweets.

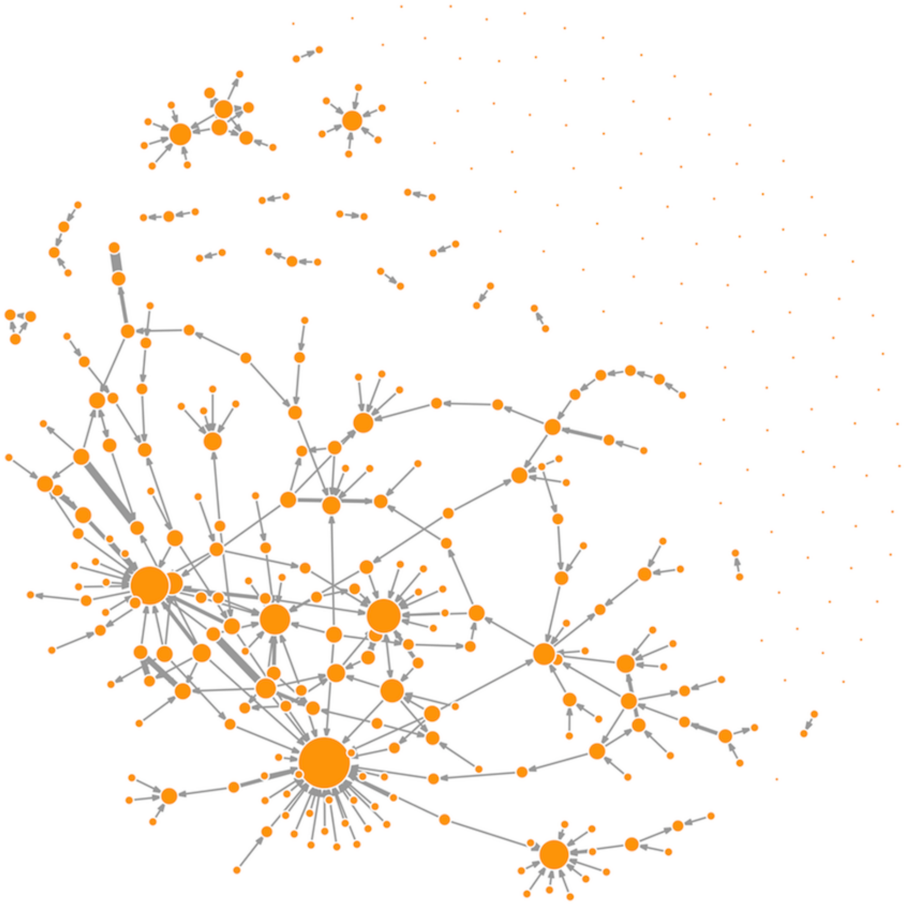


Table 3. The GLM effects of predictors on individuals' opinion leadership.

Predictors	Local opinion leadership		Global opinion leadership			
	Indegree		In-closeness		Betweenness	
	Coefficients	Standard error	Coefficients	Standard error	Coefficients	Standard error
Intercept	1.57[a]	0.31	0.000002[a]	0.000000003	−53.76	29.67
Personal attributes						
Activeness on Weibo	0.001[a]	0.0003	0.00000000006[a]	0.000000000003	0.89[a]	0.03
Verification	−0.33[a]	0.05	−0.0000000007	0.0000000005	−17.36[a]	4.98
Location	0.02	0.01	0.0000000005[a]	0.0000000001	3.72[a]	1.05
Competence						
Medical-focused	1.14[a]	0.30	0.000000001	0.000000003	−8.02	29.35
Social positions						
The number of followers	0.01[a]	0.00008	0.00000000002[a]	0.0000000000008	0.02[b]	0.01
The number of followings	−0.003	0.002	0.0000000002[a]	0.00000000002	0.34[c]	0.16
Δ R^2	34.4%		5.2%		4.2%	
Control variable						
Gender (female=1, male=0)	−0.14[a]	0.03	−0.000000002[a]	0.0000000003	7.60[c]	3.16
Δ R^2	0.0%		0.0%		0.0%	
Total R^2	34.4%		5.2%		4.2%	

[a] *P*<.001.

[b] *P* ≤.01.

[c] *P*<.05.

The Predictors of Opinion Leadership

The GLM results are reported in Table 3. The results showed that all predictors accounted for 34.4% of the variance in indegree, 5.2% of the variance in in-closeness, and 4.2% of the variance in betweenness. Personal attributes, professional knowledge, and network positions significantly affected the number retweets one node received (ie, local opinion leadership). Nevertheless, in terms of global opinion leadership, only personal attributes and network positions were significant predictors. Professional knowledge did not significantly affect this type of opinion leadership.

Personal Attributes

H1 predicted a positive effect of users' activeness on Weibo on (1) local and (2) global opinion leadership on the organ donation topic. The analysis revealed that the number of messages one posted on Weibo was significantly and positively associated with one's local opinion leadership, *B*=.001, *P*<.001, as well as global opinion leadership: *B*=6e-11, *P*<.001 for in-closeness and *B*=.89, *P*<.001 for betweenness. Hence, the data were consistent with H1 (1) and H1 (2).

H2 proposed that, compared with an unverified user, a verified user exhibits more (1) local and (2) global opinion leadership within the retweet network about organ donation. However, the results showed an opposite direction of effect. The unverified users exhibited significantly more local opinion leadership than verified users within the retweet network, *B*=−.33, *P*<.001. With regard to the global opinion leadership, they obtained a significantly higher value of betweenness than verified users, *B*=−17.36, *P*<.001. The results of the other indicator, in-closeness, did not reach significance. Hence, the unverified users showed significantly more opinion leadership than did verified users, and the data were inconsistent with H2 (1) and H2 (2).

H3 predicted positive effects of ICT development level in users' location on his or her (1) local and (2) global opinion leadership in the retweet network of organ donation message. The results showed that level of ICT development was not significantly associated with local opinion leadership. However, it was positively associated with two global opinion leadership indicators: in-closeness, *B*=5e-10, *P*<.001 and betweenness, *B*=3.72, *P*<.001. Thus, the data were inconsistent with H3 (1) but consistent with H3 (2).

Competence

H4 made predictions about the effects of professional, medical knowledge on (1) local and (2) global opinion leadership in the retweet network of organ donation messages. The results showed that medical-focused users significantly exhibited more local opinion leadership than other users, *B*=1.14, *P*<.001. However, such effect on global opinion leadership did not reach significance. Thus, the data were consistent with H4 (1) but inconsistent with H4 (2).

Social Positions

H5 and H6 considered the effects of social position on Weibo on one's opinion leadership on the organ donation topic. H5 predicted a positive effect of the number of followers on (1) local and (2) global opinion leadership. The results showed that users with a higher number of followers were more likely to exhibit more local as well as global opinion leadership in retweet network about organ donation: B=.01, P<.001 for indegree, B=2e-11, P<.001 for in-closeness, and B=.02, P=.01 for betweenness. Hence, the data were consistent with H5 (1) as well as H5 (2). H6 anticipated a positive effect of the number of followings on (1) local and (2) global opinion leadership. The analysis revealed that users with a higher number of followings showed more global opinion leadership: B=2e-10, P<.001 for in-closeness and B=.34, P=.03 for betweenness. However, it did not significantly affect local opinion leadership. Therefore, the data were inconsistent with H6 (1) but consistent with H6 (2).

Discussion

Major Findings and Implications

This study investigates organ donation information on Weibo by mapping its sharing (ie, retweet) network and examining the local as well as global opinion leadership in the network. This work explores the role of personal attributes, professional knowledge, and social position in obtaining influence according to Katz's [11] treatise. The findings reveal that all three factors predict individuals' local opinion leadership in the retweet network. Alternatively, personal attributes and social position, but not professional knowledge, are significantly associated with global opinion leadership. This study's findings significantly improve the understanding of organ donation information on social media and will be instrumental in the design of organ donation promotions on social media.

The sharing network of popular organ donation messages on Weibo is extremely sparse and centralized, resembling a star-like network structure. Only a very small portion of users in this network receives retweets from others, whereas more than 90% of users do not receive any retweets from others and occupy peripheral positions in the network. This result indicates that few central users control the flow of organ donation information and could act as critical peer leaders in organ donation promotions on Weibo. After mapping the network, subsequent analysis explores how individual and social factors affect these users' ability to influence the information flow (ie, opinion leadership). In addition, the opinion leadership on social media is conceptualized as a two-dimensional construct, including a direct influence in neighborhood (ie, local opinion leadership), as well as an indirect impact in the whole environment (ie, global opinion leadership).

The findings show that two personal attributes are significant predictors of both local and global opinion leadership on organ donation: activeness on Weibo and verification status. In detail, compared with inactive users, active users are more likely to show greater local and global opinion leadership in the organ donation information diffusion network on Weibo. This finding is consistent with theories on developing influence [30,32], as well as previous empirical research on Weibo [33,34]. Hence, the activeness of users could be a direct and simple criterion of selecting peer leaders for SNS organ donation promotions. These users exert strong influence on their neighbors and spread the information throughout the entire network and target audience.

Nevertheless, the other personal attribute, verification status, negatively impacts opinion leadership, which is the opposite of H2's prediction. This study found that, compared with verified users, unverified users are more likely to show greater local as well as global opinion leadership about organ donation on Weibo. One possible explanation is that a user's influence on social media is topic-sensitive [45]. Although a previous study claimed that verified users are perceived to be more credible than unverified ones and that their posts on Weibo received more retweets than others' posts, this study was not topic-specific and contained nine different topics from personal interests to political news [33]. The current study, however, focuses on a specific topic—organ donation. It is possible that the role of verification status varies for different topics and that Weibo users would turn to unverified rather than verified users for opinions about organ donation.

The other possible explanation could be that, in general, verified users may enjoy less rather than more credibility than unverified ones. Indeed, the Chinese government has hired a large number of people to fabricate posts on popular websites and social media, and the number of pseudonymous and deceptive social media posts could reach 488 million a year [46]. Although the government has never publicly or officially admitted such an operation, the general public in China is fully aware of it. Thus, the verification badge on Weibo could backfire. People may regard users who have received official verification as government employees as individuals who may intentionally manipulate public opinion on Weibo. This could be why unverified users exhibit greater local and global opinion leadership about organ donation than verified users on Weibo. However, both explanations need further examination. Health communication professionals should be aware of this counterintuitive finding and be more cautious when choosing verified users on Weibo as peer leaders to promote organ donation.

The third personal attribute examined in this study is the level of ICT development in one's location. Unlike the abovementioned two attributes, which are relevant to a user's activities and identity on Weibo, this one is an environmental factor. The results show that a person's direct impact on the neighborhood (ie, local opinion leadership) is highly associated with his or her characteristics and identity on Weibo but not with the ICT development level in his or her area. A user's indirect influence on other users (ie, global opinion leadership) depends on that user's characteristics on Weibo as well as this environmental factor. Indeed, a previous study found that ICT development was highly associated with users' influence in friendship networks on social media [36]. This study extends the previous research on social (ie, following-follower) networks to an information network on Weibo. The results show that, with regard to the organ donation topic on Weibo, although the ICT development level in one's local network does not affect his or her direct influence on others, it significantly impacts the user's indirect influence and the ability to control information flows in the whole environment. Therefore, users from ICT mature areas can be recruited as peer leaders for SNS organ donation promotion targeting on a wide range of audience groups.

The second type of opinion leadership predictor suggested by Katz [11] is professional knowledge on the topic. This study shows that users with medical knowledge exhibit significantly greater local opinion leadership about organ donation on Weibo than users without such knowledge. However, this effect is not significant for global opinion leadership. It is possible that a user's professional knowledge on a certain topic is critical for obtaining opinion leadership when the information flows follow the "two-step flow." Hence, as the results reveal, it significantly affects local opinion leadership on the organ donation topic. However, competence or credibility may become less important when individuals want to wield global influence over the retweet network, within which multiple avenues of information flow coexist. Thus, although medical professionals are influential organ donation opinion leaders, their impact is limited to their close neighbors.

The last predictor of opinion leadership included in this study is a user's social position on Weibo. The results reveal that compared with obtaining local opinion leadership, securing global influence requires a well-connected social location in the network. For local opinion leadership, the number of followers, but not followings, is a significant predictor. The number of a user's followers is the number of users on Weibo who will be directly exposed to his or her posts (ie, the user's direct audience). As documented in these results, the larger the size of a user's direct audience, the greater level of that user's local influence. On the other hand, global opinion leadership depends on not only the size of one's direct audience but also the size of the user's information sources (ie, followings) on Weibo. Users with large audiences and many information sources occupy well-connected positions in the network and have updated information on the topic, thus exhibiting more global opinion leadership than others who occupy peripheral social positions. For public health professionals, they may recruit peer leaders according to campaign objectives. A user with a large direct audience will be competent to impact his or her neighbors, but only users with a large audience and many information sources will be capable of controlling the dissemination of organ donation information on Weibo.

Although previous organ donation campaigns have employed social media, the campaign advertisements and strategies were specifically designed for college students, and the information dissemination was mainly controlled by the researchers [5,6]. This procedure, however, may have limited generalizability to other populations and also to other contexts outside the confines of a controlled environment [47]. To explore opinion leadership on social media in a natural context, this study analyzes the general public's organ donation discussion and retweeting behavior on a popular Chinese social networking site. This unobtrusive approach offers implications for recruiting peer leaders on social media to promote organ donation. In sum, the findings of this study indicate that health campaign designers may recruit peer leaders in SNS organ donation promotions to facilitate information sharing among target audience. Users who are unverified, active, well connected, and experienced with ICT will accelerate the sharing of organ donation messages in the global environment. Medical professionals such as organ transplant surgeons who can wield a great amount of influence on their direct connections could also effectively participate in promoting organ donation on social media.

Limitations and Future Research

There are several areas worthy of further research in opinion leadership in the topic of organ donation on social media. First, examining the retweet paths of all the organ donation messages (n=6701) would yield an extremely large dataset and be computationally intensive, so this study focuses on only the most popular messages. In fact, this study initiates an exploration of opinion leadership of organ donation promotion on social media with an innovative and advanced method. Future research may replicate this research on other SNS platforms or with a larger dataset. Second, this study analyzes a snapshot of the retweet network instead of a dynamic diffusion network that evolves over time. Subsequent work may employ more sophisticated data mining and data analyzing techniques to detect how organ donation messages go viral on social media and who facilitates the dissemination, which would offer valuable information for future SNS organ donation promotion. Third, apart from medical knowledge, some other factors such as experience with organ donation may contribute to a user's opinion leadership on the topic of organ donation. Future studies could explore other measures or indicators of users' competence on the topic of organ donation. Fourth, this study examines the retweet network of all organ donation tweets regardless of their content. However, the structure of the retweet network may vary by how organ donation is covered or framed in the tweets. Future research should investigate whether content shapes the retweet paths and opinion leadership. For example, are myths about organ donation disseminated the same way as stories about an organ recipient? Findings from this research will greatly enhance the design and implementation of organ donation campaigns using social media.

Acknowledgments

The authors thank Dr Tai-Quan Peng and Mr Xiaohui Wang for their helpful comments on an earlier draft of this paper.

Conflicts of Interest

None declared.

References

1. irodat. 2015. IRODaT Newsletter 2015 URL: http://www.irodat.org/img/database/pdf/NEWSLETTER2016_FirstEdition.pdf [accessed 2017-09-08] [WebCite Cache ID 6tJjTOKEX]
2. Huang J, Wang H, Fan ST, Zhao B, Zhang Z, Hao L, et al. The national program for deceased organ donation in China. Transplantation 2013 Jul 15;96(1):5-9. [doi: 10.1097/TP.0b013e3182985491] [Medline: 23743728]
3. China Central Television (CCTV). news.sina. 1.5 million patients need organ transplant per year, but less than 10,000 of them are able to received one URL: http://news.sina.com.cn/c/sd/2010-11-15/002521469883.shtml [accessed 2017-09-08] [WebCite Cache ID 6tJjpAgO2]
4. china-organdonation. 2017. Statistics URL: http://www.china-organdonation.org/ [accessed 2017-09-08] [WebCite Cache ID 6tJk3cNrl]
5. Stefanone M, Anker AE, Evans M, Feeley TH. Click to “like” organ donation: the use of online media to promote organ donor registration. Prog Transplant 2012 Jun;22(2):168-174. [doi: 10.7182/pit2012931] [Medline: 22878074]
6. Smith SW, Hitt R, Park HS, Walther J, Liang YJ, Hsieh G. An effort to increase organ donor registration through intergroup competition and electronic word of mouth. J Health Commun 2016 Jan;21(3):376-386. [doi: 10.1080/10810730.2015.1095815] [Medline: 26735448]
7. Cucchetti A, Zanello M, Bigonzi E, Pellegrini S, Cescon M, Ercolani G, et al. The use of social networking to explore knowledge and attitudes toward organ donation in Italy. Minerva Anestesiol 2012 Oct;78(10):1109-1116 [FREE Full text] [Medline: 23059515]
8. expandedramblings. By the number: 59 amazing Weibo statistics (December 2016) URL: http://expandedramblings.com/index.php/weibo-user-statistics/ [accessed 2017-03-01] [WebCite Cache ID 6odN1lrdE]
9. Shi J. repository.ntu. Singapore: Nanyang Technological University; 2017. A multi-method approach to the study of “attitude” in organ donation research URL: https://repository.ntu.edu.sg/handle/10356/70764 [accessed 2017-12-20] [WebCite Cache ID 6vqTiwiBH]
10. Valente TW, Pumpuang P. Identifying opinion leaders to promote behavior change. Health Educ Behav 2007 Dec;34(6):881-896. [doi: 10.1177/1090198106297855] [Medline: 17602096]
11. Katz E. The two-step flow of communication: an up-to-date report on an hypothesis. Public Opin Q 1957;21:61-78. [doi: 10.1086/266687]
12. Katz E, Lazarsfeld P. Personal Influence. New York, NY: Free Press; 1955.
13. Rogers E, Cartano D. Methods of measuring opinion leadership. Public Opin Q 1962;26(3):435-441. [doi: 10.1086/267118]
14. Boster F, Kotowski M, Andrews K, Serota K. Identifying influence: development and validation of the connectivity, persuasiveness, and maven scales. J Commun 2011;61(1):178-196. [doi: 10.1111/j.1460-2466.2010.01531.x]
15. boyd D, Ellison N. Social network sites: definition, history, and scholarship. J Comput Mediat Commun 2007;13(1):210-230. [doi: 10.1111/j.1083-6101.2007.00393.x]
16. Xu WW, Sang Y, Blasiola S, Park HW. Predicting opinion leaders in Twitter activism networks: the case of the Wisconsin recall election. Am Behav Sci 2014 Mar 13;58(10):1278-1293. [doi: 10.1177/0002764214527091]
17. Hilbert M, Vasquez J, Halpern D, Valenzuela S, Arriagada E. One step, two step, network step? Complementary perspectives on communication flows in Twittered citizen protests. Soc Sci Comput Rev 2017;35(4):444-461. [doi: 10.1177/0894439316639561]
18. Park CS. Does Twitter motivate involvement in politics? Tweeting, opinion leadership, and political engagement. Comput Human Behav 2013 Jul;29(4):1641-1648. [doi: 10.1016/j.chb.2013.01.044]
19. Bodendorf F, Kaiser C. Detecting opinion leaders and trends in online social networks. New York, NY: ACM; 2009 Presented at: The 2nd ACM workshop on social web search and mining; November 02-02; Hong Kong, China p. 65-68. [doi: 10.1145/1651437.1651448]
20. Valente T. Social networks and health. New York, NY: Oxford University Press; 2010.
21. Weimann G, Tustin DH, van Vuuren D, Joubert JPR. Looking for opinion leaders: traditional vs. modern measures in traditional societies. Int J Public Opin Res 2007 Mar 04;19(2):173-190. [doi: 10.1093/ijpor/edm005]
22. Kratzer J, Lettl C. Distinctive roles of lead users and opinion leaders in the social networks of schoolchildren. J Consum Res 2009 Dec 01;36(4):646-659. [doi: 10.1086/599324]
23. Lee SH, Cotte J, Noseworthy TJ. The role of network centrality in the flow of consumer influence. J Consum Psychol 2010 Jan;20(1):66-77. [doi: 10.1016/j.jcps.2009.10.001]
24. Stewart FA. A Sociometric Study of Influence in Southtown. Sociometry 1947 Feb;10(1):11. [doi: 10.2307/2785558]
25. Hanneman RA, Riddle M. faculty.ucr.edu. 2005. Introduction to social network methods URL: http://faculty.ucr.edu/~hanneman/nettext/ [accessed 2017-12-20] [WebCite Cache ID 6vqW7BmKH]
26. Freeman L. Centrality in social networks conceptual clarification. Soc Networks 1978;1(3):215-239. [doi: 10.1016/0378-8733(78)90021-7]
27. Borgatti S, Everett M, Johnson J. Analyzing Social Networks. Thousand Oaks, CA: Sage; 2013.
28. Burt RS. The network structure of social capital. Res Organ Behav 2000;22:345-423. [doi: 10.1016/S0191-3085(00)22009-1]
29. Burt R. pdfs.semanticscholar. 2001. The social capital of structural holes URL: https://pdfs.semanticscholar.org/fc11/32dbf80dd11364a7d5bbb19d6d5880ce8e51.pdf [accessed 2017-12-20] [WebCite Cache ID 6vqX2cZE1]

30. Huffaker D. Dimensions of leadership and social influence in online communities. Hum Commun Res 2010;36(4):593-617. [doi: 10.1111/j.1468-2958.2010.01390.x]
31. Rice RE. New patterns of social structure in an information society. In: Schement J, Lievrouw L, editors. Competing visions, complex realities: Social aspects of the information society. Norwood, NJ: Ablex; 1987:107-120.
32. Weimann G. The influentials. Albany, NY: The State University of New York Press; 1994.
33. Zhang L, Peng T, Zhang Y, Wang X, Zhu JJ. Content or context: which matters more in information processing on microblogging sites. Comput Human Behav 2014 Feb;31:242-249. [doi: 10.1016/j.chb.2013.10.031]
34. Shi J, Wang X, Peng T, Chen L. Understanding interactions in virtual HIV communities: a social network analysis approach. AIDS Care 2017 Feb;29(2):239-243. [doi: 10.1080/09540121.2016.1210074] [Medline: 27433828]
35. Grabowicz P, Babaei M, Kulshrestha J, Weber I. The road to popularity: The dilution of growing audience on twitter. 2016 Presented at: The tenth international AAAI conference on web and social media (ICWSM 2016); May 17–20; Cologne, Germany URL: http://www.aaai.org/ocs/index.php/ICWSM/ICWSM16/paper/download/13134/12790
36. Wang X, Shi J, Chen L, Peng T. An examination of users' influence in online HIV/AIDS communities. Cyberpsychol Behav Soc Netw 2016 May;19(5):314-320. [doi: 10.1089/cyber.2015.0539] [Medline: 27186897]
37. DiMaggio P, Hargittai E, Neuman W. Social implications of the Internet. Annu Rev Sociol 2001;27:307-336 [FREE Full text]
38. Eastin MS, LaRose R. Internet self-efficacy and the psychology of the digital divide. J Comput Mediat Commun 2000;6(1):-. [doi: 10.1111/j.1083-6101.2000.tb00110.x]
39. Lyons B, Henderson K. Opinion leadership in a computer-mediated environment. J Consum Behav 2005 Sep;4(5):319-329. [doi: 10.1002/cb.22]
40. Lima CX, Lima MV, Cerqueira RG, Cerqueira TG, Ramos TS, Nascimento M, et al. Organ donation: cross-sectional survey of knowledge and personal views of Brazilian medical students and physicians. Transplant Proc 2010 Jun;42(5):1466-1471. [doi: 10.1016/j.transproceed.2009.11.055] [Medline: 20620455]
41. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960;20(1):37-46. [doi: 10.1177/001316446002000104]
42. Csardi G, Nepusz T. Necsi.edu. 2006. The igraph software package for complex network research URL: http://www.necsi.edu/events/iccs6/papers/c1602a3c126ba822d0bc4293371c.pdf [accessed 2017-12-20] [WebCite Cache ID 6vqY9Whg9]
43. Song Z, Liu W, Ma L, Dunford M. Measuring spatial differences of informatization in China. Chin Geogr Sci 2013 Nov 11;24(6):717-731. [doi: 10.1007/s11769-013-0646-1]
44. International Telecommunication Union. itu.int. Digital access index (DAI) URL: http://www.itu.int/ITU-D/ict/dai/ [accessed 2017-12-20] [WebCite Cache ID 6vqZ9u0OY]
45. Weng J, Lim E, Jiang J, He Q. TwitterRank: Finding topic-sensitive influential Twitterers. 2010 Presented at: The third ACM international conference on web search & data mining; February 03-06; New York, NY URL: http://ink.library.smu.edu.sg/cgi/viewcontent.cgi?article=1503&context=sis%5Fresearch
46. King G, Pan J, Roberts M. How the Chinese government fabricates social media posts for strategic distraction, not engaged argument. Am Polit Sci Rev 2017;111(3):484-501. [doi: 10.1017/S0003055417000144]
47. Shi J, Poorisat T, Salmon CT. The use of social networking sites (SNSs) in health communication campaigns: review and recommendations. Health Commun 2018;33(1):49-56. [doi: 10.1080/10410236.2016.1242035] [Medline: 27858464]

Abbreviations

GLM: general linear model
ICT: information and communications technology
SD: standard deviation
SNS: social networking site

Edited by G Eysenbach; submitted 05.03.17; peer-reviewed by T Feeley, R Hitt; comments to author 15.07.17; revised version received 26.09.17; accepted 16.11.17; published 09.01.18.

Please cite as:
Shi J, Salmon CT
Identifying Opinion Leaders to Promote Organ Donation on Social Media: Network Study
J Med Internet Res 2018;20(1):e7
URL: http://www.jmir.org/2018/1/e7/
doi:10.2196/jmir.7643
PMID:29317384

Original Paper

Efficacy of a Web-Based Safety Decision Aid for Women Experiencing Intimate Partner Violence: Randomized Controlled Trial

Jane Koziol-McLain[1], RN, PhD; Alain C Vandal[2,3], PhD; Denise Wilson[4], RN, PhD; Shyamala Nada-Raja[5], PhD; Terry Dobbs[1], MA; Christine McLean[1], PGDipCounsTh; Rose Sisk[2], BSc (Hons); Karen B Eden[6], PhD; Nancy E Glass[7], RN, PhD

[1]Centre for Interdisciplinary Trauma Research, Faculty of Health and Environmental Sciences, Auckland University of Technology, Auckland, New Zealand

[2]Department of Biostatistics and Epidemiology, Faculty of Health and Environmental Sciences, Auckland University of Technology, Auckland, New Zealand

[3]Ko Awatea, Counties Manukau Health, Auckland, New Zealand

[4]Taupua Waiora Centre for Māori Health Research, Faculty of Health and Environmental Sciences, Auckland University of Technology, Auckland, New Zealand

[5]Department of Preventive and Social Medicine, University of Otago, Dunedin, New Zealand

[6]Department of Medical Informatics and Clinical Epidemiology, School of Medicine, Oregon Health Sciences University, Portland, OR, United States

[7]Johns Hopkins Center for Global Health, Johns Hopkins University, Baltimore, MD, United States

Corresponding Author:
Jane Koziol-McLain, RN, PhD
Centre for Interdisciplinary Trauma Research
Faculty of Health and Environmental Sciences
Auckland University of Technology
Private Bag 92006
Auckland, 1142
New Zealand
Phone: 64 9 921 9670
Fax: 64 9 921 9796
Email: jane.koziol-mclain@aut.ac.nz

Abstract

Background: Intimate partner violence (IPV) is a human rights violation and leading health burden for women. Safety planning is a hallmark of specialist family violence intervention, yet only a small proportion of women access formal services. A Web-based safety decision aid may reach a wide audience of women experiencing IPV and offer the opportunity to prioritize and plan for safety for themselves and their families.

Objective: The aim of this study was to test the efficacy of a Web-based safety decision aid (*isafe*) for women experiencing IPV.

Methods: We conducted a fully automated Web-based two-arm parallel randomized controlled trial (RCT) in a general population of New Zealand women who had experienced IPV in the past 6 months. Computer-generated randomization was based on a minimization scheme with stratification by severity of violence and children. Women were randomly assigned to the password-protected intervention website (safety priority setting, danger assessment, and tailored action plan components) or control website (standard, nonindividualized information). Primary endpoints were self-reported mental health (Center for Epidemiologic Studies Depression Scale-Revised, CESD-R) and IPV exposure (Severity of Violence Against Women Scale, SVAWS) at 12-month follow-up. Analyses were by intention to treat.

Results: Women were recruited from September 2012 to September 2014. Participants were aged between 16 and 60 years, 27% (111/412) self-identified as Māori (indigenous New Zealand), and 51% (210/412) reported at baseline that they were unsure of their future plans for their partner relationship. Among the 412 women recruited, retention at 12 months was 87%. The adjusted estimated intervention effect for SVAWS was −12.44 (95% CI −23.35 to −1.54) for Māori and 0.76 (95% CI −5.57 to 7.09) for non-Māori. The adjusted intervention effect for CESD-R was −7.75 (95% CI −15.57 to 0.07) for Māori and 1.36 (−3.16 to 5.88) for non-Māori. No study-related adverse events were reported.

Conclusions: The interactive, individualized Web-based *isafe* decision aid was effective in reducing IPV exposure limited to indigenous Māori women. Discovery of a treatment effect in a population group that experiences significant health disparities is a welcome, important finding.

Trial Registration: Australian New Zealand Clinical Trials Registry (ANZCTR): ACTRN12612000708853; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?ACTRN=12612000708853 (Archived by Webcite at http://www.webcitation/61MGuVXdK)

(J Med Internet Res 2018;19(12):e426) doi:10.2196/jmir.8617

KEYWORDS

eHealth; intimate partner violence; randomized controlled trial; New Zealand; depression; population groups

Introduction

Intimate partner violence (IPV) against women is a human rights violation with significant health consequences [1,2]. There is a substantial body of literature documenting the prevalence of violence against women; international evidence suggests that one in three women have experienced IPV or sexual violence [2]. The impacts of IPV on multiple aspects of health have been extensively documented—including mental health, sexual and reproductive health, and chronic conditions—and represent a significant health burden for women [3]. The World Health Organization Global Plan of Action calls for strengthening the role of the health system in addressing violence against women and girls [4,5], yet, there is a paucity of evidence testing theoretically informed interventions for women. In a Cochrane systematic review evaluating advocacy interventions providing safety planning or facilitating access to community IPV resources, 8 of the 13 studies recruited women from health care settings [6]. Although there were some benefits from brief advocacy interventions (ie, may provide short-term reduction in anxiety, distress, depression, and violence recurrence), there was significant heterogeneity among the studies leading to “uncertainty about the magnitude of benefit and the impact of abuse severity and the setting.” In addition, a minority of women experiencing IPV access formal IPV services. For example, among women participants of the New Zealand Violence Against Women Study who had ever been physically or sexually abused by a partner, 69% (658/956) had never told a formal resource about their partner’s behavior [7].

Electronic health (eHealth) provides an opportunity to reach a broad population of women and deliver an interactive, tailored intervention at no cost, at any time of day, and free from the stigma that may be associated with face-to-face interventions. A novel Web-based safety decision aid for women experiencing IPV was recently developed in the United States [8]. The decision aid, informed by an empowerment model [9] and decision-aid science [10], includes priority-setting activities, risk assessment with feedback, and tailored safety planning. In a randomized controlled trial (RCT), women’s decisional conflict reduced after a single use of the eHealth intervention [11]. To build further evidence, we conducted a concurrent replication of the US trial regionalized for Aotearoa New Zealand culture [12]. The New Zealand (*isafe*) trial advances the US trial in providing fully automated Web-based trial recruitment, eligibility screening, and consent. In this RCT, we tested the efficacy of an interactive Web-based safety decision aid. We hypothesized that the decision aid would improve mental health and reduce IPV exposure.

Methods

Trial Design and Participants

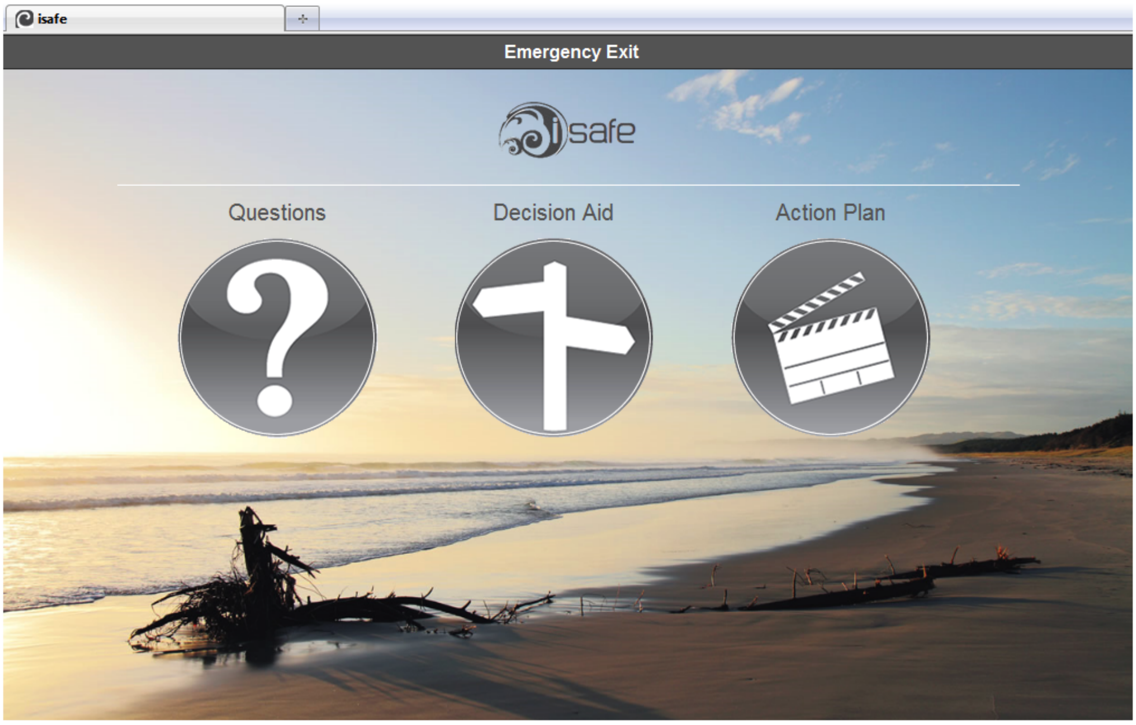
Our study protocol is described elsewhere [13]. Briefly, the *isafe* study consisted of a two-arm parallel RCT of a Web-based safety decision aid for women experiencing IPV. Efficacy of the decision aid was assessed primarily using participants’ exposure to IPV and mental health after 12 months of access to the decision aid. Participation was open to English-speaking women aged 16 years or older in New Zealand who reported exposure to current IPV. They also needed access to a safe email address (to which only they had password-protected access) to send and receive study-related information. The trial was approved by the Auckland University of Technology Ethics Committee (reference number 12/51).

Complete details regarding participant recruitment and engagement can be found elsewhere [14]. Participants enrolled in the study by accessing a secure New Zealand registration website; the most common referral source was a Web classified advertisement. A participant was considered enrolled once they had visited the registration website and (1) met eligibility criteria, (2) consented to participate, (3) provided contact information, and (4) passed a validation check. The validation check to reduce the risk of fraudulent enrollment involved automated matching against the New Zealand electoral roll (based on name and address) or a manual process of logic checking against information available from Web searching.

Randomization and Masking

Once enrolled, participants were randomized by software to the control or intervention arm. Randomization was based on a minimization scheme. Two stratification factors (severity of violence and children) and two random factors each with two equiprobable levels were used to achieve the minimization. Severity of violence factor was dichotomous, based on one positive response to the IPV eligibility items versus two or more. The children factor was also dichotomous, identifying whether the participant had one or more children versus none.

Figure 1. Screenshot of isafe navigation.



After randomization, participants were sent an email with a username, password, and URL access for the website. The participant's allocation was kept secret from herself and the study team in New Zealand. There was no procedural difference between arms until after baseline measures were obtained. After baseline measures collection, participants in different arms were exposed to different screen contents, corresponding to the control and active interventions described below (see eg, Figure 1).

Procedures

Women randomly assigned to the intervention group were able to access a Web-based safety decision aid throughout the 1-year postbaseline follow-up period via the secure password-protected trial website. The decision aid intervention included three components. The first component was a safety priority setting activity based on a multicriteria decision model developed in the United States [8]. The five criteria (priorities) were evaluated for the New Zealand context [12] and minor wording changes made to the descriptions [13]. Women moved a sliding bar toward the priority that was most important to them for all pairwise combinations. Through a series of matrix computations using the analytic hierarchy process [15], the program provided feedback to the participant summarizing her priorities. The second component was the Danger Assessment (DA) [16] or Danger Assessment-Revised (DA-R; for female same-sex relationships) [17]. Women completed the DA or DA-R and received immediate scored feedback on their level of danger for severe or lethal violence in the intimate relationship, ranging from variable danger to extreme danger. The third and final component was an interactive process using an underlying matrix of resources to help women develop an individually tailored action plan. The matrix included local, regional, and national resources; and tips about safety for the participant and her children based on her safety priorities and DA or DA-R scores.

Women randomly assigned to the control group were able to access a standardized list of resources and a standardized emergency safety plan throughout the 1-year postbaseline follow-up period via the secure password-protected trial website. They were not provided with individualized feedback or tailored action plan.

Participants were followed up for 1 year following completion of their baseline measures. Assessments were scheduled to take place at baseline, then 3, 6, and 12 months after completion of baseline. Women were invited to complete assessments at each time period regardless of prior assessment completions. Retention mechanisms included automated emails sent to participants at 1, 2, 4, 8, and 10 months. If a participant was 6 weeks late for a scheduled follow-up visit, the research team sent email or phone reminders (following the woman's safety instructions).

Logs were maintained by study personnel documenting phone and email contacts with participants, potential unanticipated events, and other issues affecting the trial (such as server interruptions). Study safety protocols addressed computer safety procedures; confidentiality measures; support and referral procedures in case of violence or distress, including suicidality; responding to calls to the study site by partners or children; and investigation and reporting of unanticipated events. In addition to safety review during data monitoring committee (DMC) meetings, study logs and media reports of serious IPV events were monitored during the trial. The study protocol outlined the safety review process, which for study-related serious adverse events would involve escalation to the principal investigator and chairs of the ethics and data monitoring committees and reporting to the funder and university legal counsel.

Outcomes

The mental health primary outcome consisted of depression measured by the Center for Epidemiologic Studies Depression Scale-Revised (CESD-R) [18-20] at 12 months. The CESD-R score, ranging from 0 to 80, was used. Higher scores are associated with more depression symptomatology, with scores of at least 16 indicative of depression. The IPV exposure primary outcome consisted of the Severity of Violence Against Women Scale (SVAWS) [21] at 12 months. For the SVAWS we used a 1 (never) to 4 (many times) subjective frequency scale for all items; total SVAWS score ranges from 46 to 196, with higher scores associated with greater exposure to IPV. Secondary violence domain outcomes included SVAWS total score at 3 and 6 months, SVAWS subscales (threats of violence, acts of violence, and sexual violence), and Women's Experience with Battering (WEB) [22]. Secondary mental health domain outcomes included CESD-R score at 3 and 6 months, Post-Traumatic Stress Disorder Checklist-Civilian Version [23,24], Alcohol Use Disorder Identification Test dichotomized [25,26], and Drug Abuse Screening Tool [27,28]. Decisional process secondary outcome was measured using the Decisional Conflict Scale [29], safety-seeking behavior using the Safety Checklist [30,31], and Safety Checklist Helpfulness. Assessment schedules and time points are included in Multimedia Appendix 1.

Statistical Analysis

As described in the published protocol [13], the figure of 340 women accounts for an upper limit of 35% dropout by 12 months based on attrition rates for previous Web-based studies (eg, the Recovery via Internet from Depression trial). Results from a New Zealand trial support the planned study to have 80% power to detect a 37% reduction in depression (CESD-R). Data from an international IPV study indicated that the proposed study would have power superior to 80% to detect a difference of −11.2 in IPV exposure (corresponding to ≈ ½ standard deviation baseline Severity Violence Against Women Scale).

Plans for all inferential analyses were finalized before allocation unblinding in a full statistical analysis plan. Although randomization was technically carried out before baseline assessment, the allocation remained entirely concealed until all baseline information was obtained. Accordingly, the intention-to-treat (ITT) analysis set consisted of all participants who provided data at baseline on a primary outcome. The main analyses were carried out in a modified ITT analysis set, consisting of the ITT analysis set from which participants with no postbaseline assessments were excluded. Sensitivity analyses were carried out in the ITT set to assess the effect of this exclusion on treatment effect estimates [32]. A per-protocol analysis set was also defined, excluding from the modified ITT set any participant identified with an eligibility violation (such as from repeat enrollment) or a major protocol violation (not receiving the intervention) and assigning the participant to the allocation arm corresponding to her actual uptake of the intervention.

All postbaseline assessments for a given outcome were entered in the regression models, initially assumed to be normal with a participant-specific random effect, adjusted for the baseline value, other covariates as described below, and placing the intervention in interaction with the assessment time, taken as a factor. A blind review, absent any information regarding allocation, was undertaken for each outcome for an assessment of missingness, a visual assessment of residual normality, an assessment of the covariance structure of the repeated measures, and an assessment of candidate covariates (ethnicity, children, paid employment, and age group) to include in the final analysis. In a model including all covariates, assessment time but no information regarding allocation, only covariates achieving a partial coefficient of determination larger than 1% were retained to adjust the final model. Subgroup analyses regarding children in their care versus not and Māori ethnicity versus not were planned for the primary outcomes. These proceeded along the same model as above, including a three-way interaction between treatment, assessment time, and subgroup indicator. Although women could select multiple ethnicities, a priori hierarchy prioritized Māori ethnicity.

Missing data at baseline was imputed using the mode of the variable in the observed ITT values [33]. Missingness in both primary outcomes was found to be significantly related to the last observed value of the outcome in question, to the assessment time being at 12 months in the case of CESD-R but to no other baseline variable, indicating that adjusting for baseline and fitting the available repeated measures with a suitable (nondiagonal) covariance structure appropriately removed the risk of bias from missingness under a missing at random assumption [34].

The sensitivity analyses to be carried out in the full ITT set consisted in the production of point estimates only with missing values singly imputed at each assessment time according to four different schemes: (1) ITT extension: missing values because of attrition imputed as the mean of the observed values in the control arm; missing values from nondropouts imputed as the mean of the arm to which the participant was allocated; (2) Return to baseline: all missing values imputed as the baseline value; (3) Worst case for intervention: missing values imputed as the worst intervention arm outcome in the intervention arm, and the best control arm outcome in the control arm; and (4) Best case for intervention: missing values imputed as the best intervention arm outcome in the intervention arm, and the worst control arm outcome in the control arm.

Mediation analyses based on multivariate regression were to be completed only if the intervention proved to be significantly related to the hypothesized mediators at a 0.15 level. All other tests of hypotheses were carried out using a significance level of 0.05 and two-sided alternatives. There was no adjustment for multiple testing. All analyses were carried out in the R software environment, version 3.1 (R Foundation for Statistical Computing, Vienna), using packages *lme4* [35] and *nlme*. An independent DMC was formed and met for the first time in September 2013, thereafter meeting four more times at approximately 6-month intervals. The role of the DMC was detailed in a charter.

Results

Participants

The 412 participating women (Figure 2) were typically young (interquartile range [IQR]=23, 36 years) and lived in a main urban area (83%); 45% (185/412) were responsible for the care of one or more children living in the household, 40% (165/412) were in paid employment, 27% of participants self-identified as Māori compared with 12% among women 15 years of age or older living in New Zealanders (2013 census, NZ Stats table), and 30% (124/412) of participants lived in neighborhoods in the highest deprivation quintile compared with the expected 20% (82/412).

Baseline Characteristics

At baseline, partners were often living with the participant (227/412, 55%). Women's decisional conflict about their partner relationship was evident: half of the participants were "unsure" of the future of their relationships, with one-quarter planning to end the relationship and one-quarter planning to remain in the relationship. Most women accessed the *isafe* study on a computer in their home (259/412, 63%) or the home of a friend or family member (37/412, 9%). The median SVAWS score was 84 (IQR 67, 103). The median CESD-R score was 32 (IQR 18, 50). A large proportion of *isafe* participants (309/412, 75%) evidenced depressive symptoms (score greater than 16); one in 12 reported suicidal thoughts ("I wished I was dead") "nearly every day" in the prior 2 weeks.

Baseline characteristics, including violence severity and depression, were similar for women allocated to control or intervention (Table 1). The known inequities of race and socioeconomic deprivation were evident among our study participants: among the 113 Māori participants, 42.5% lived in the highest deprivation quintile; this compared with 25.1% among the 299 non-Maori participants. Although Māori were overrepresented in high deprivation neighborhoods, deprivation was not associated with group or outcome (and thus, not a confounder). Seventy-five percent of participating women completed all three follow-up assessments (73% of women assigned to control group and 76% of women assigned to intervention group). Retention rates at 3, 6, and 12 months were 81%, 83%, and 87%, respectively.

Figure 2. Trial profile (ITT=Intention to treat; PP=Per-protocol).

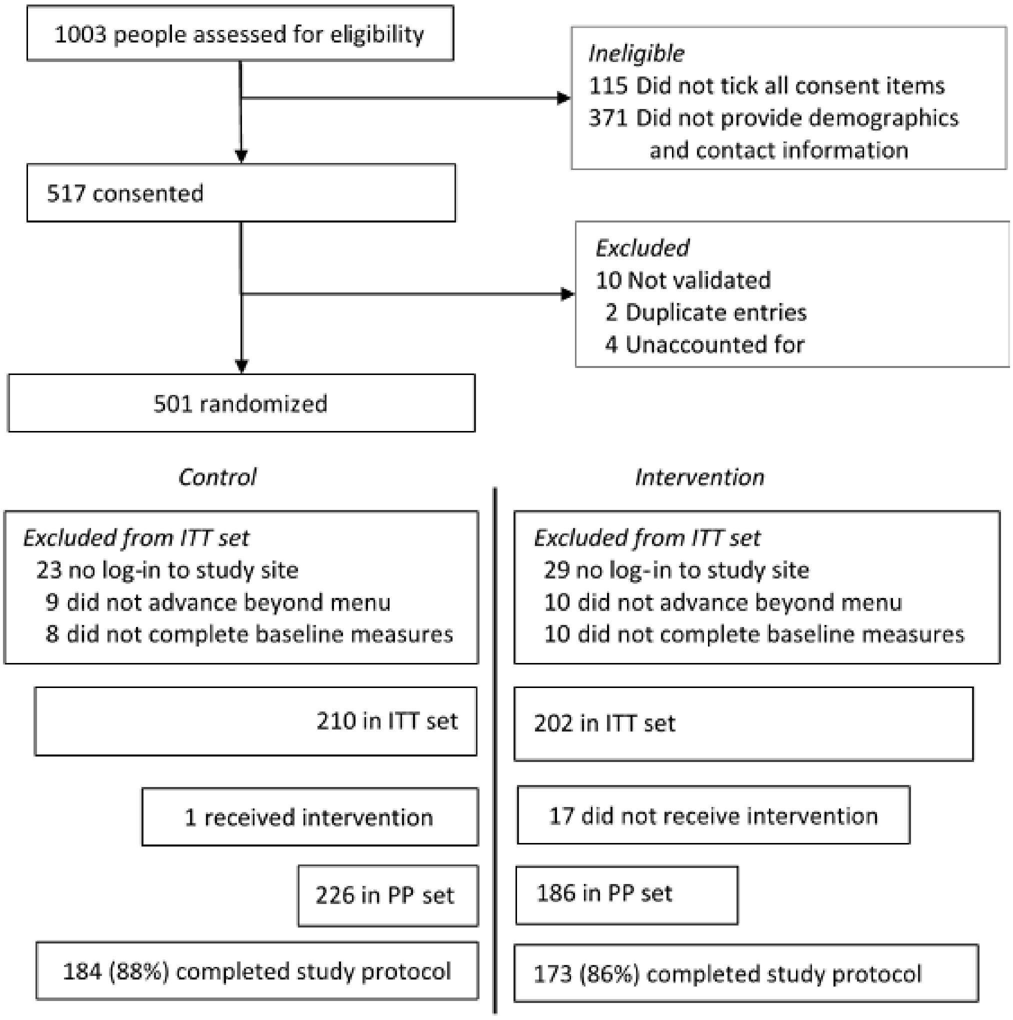


Table 1. Baseline characteristics of the intention-to-treat analysis set.

Characteristics	Control (n=210)	Intervention (n=202)	Total (N=412)
Age (years)			
Median (range)	29.0 (16-59)	29.0 (17-58)	29.0 (16-59)
IQR[a]	23.0-36.0	24.0-36.0	23.0-36.0
Ethnic group[b], n (%)			
European	149 (71.0)	148 (73.3)	297 (72.1)
Māori (indigenous New Zealand)	57 (27.1)	56 (27.7)	113 (27.4)
Asian	24 (11.4)	18 (8.9)	42 (10.2)
Pasifika	21 (10.0)	21 (10.4)	42 (10.2)
Other	3 (1.4)	4 (2.0)	7 (1.7)
None	3 (1.4)	3 (1.5)	6 (1.5)
Children[c], n (%)			
One child or more	96 (45.7)	90 (44.6)	186 (45.1)
No children	111 (52.9)	109 (54.0)	220 (53.4)
Employment status, n (%)			
In paid employment	84 (40.0)	79 (39.1)	163 (39.6)
Unemployed but looking for paid work	55 (26.2)	51 (25.2)	106 (25.7)
Unemployed, not looking for paid work	10 (4.8)	8 (4.0)	18 (4.4)
On a benefit	41 (19.5)	45 (22.3)	86 (20.9)
Other	16 (7.6)	13 (6.4)	29 (7.0)
Deprivation quintiles, n (%)			
1 (lowest deprivation)	21 (10.0)	21 (10.4)	42 (10.2)
2	31 (14.8)	31 (15.3)	62 (15.0)
3	36 (17.1)	45 (22.3)	81 (19.7)
4	56 (26.7)	41 (20.3)	97 (23.5)
5 (highest deprivation)	59 (28.1)	64 (31.7)	123 (29.9)
Unknown	7 (3.0)	0 (0.0)	7 (1.7)
Locality, n (%)			
Main urban area	174 (82.9)	169 (83.7)	343 (83.3)
Other	29 (13.8)	33 (16.3)	62 (15.0)
Partner relationship, n (%)			
Husband or wife	43 (20.5)	34 (16.8)	77 (18.7)
Ex-husband or ex-wife	1 (0.5)	5 (2.5)	6 (1.5)
Separated husband or wife	12 (5.7)	9 (4.5)	21 (5.1)
Boyfriend or girlfriend	57 (27.1)	56 (27.7)	113 (27.4)
Ex-boyfriend or ex-girlfriend	20 (9.5)	18 (8.9)	38 (9.2)
De facto partner	60 (28.6)	69 (34.2)	129 (31.3)
Ex-de facto partner	8 (3.8)	9 (4.5)	17 (4.1)
Same sex partner	5 (2.4)	1 (0.5)	6 (1.5)
Ex-same sex partner	4 (1.9)	0 (0.0)	4 (1.0)
Length of partner relationship (years)			
Median (range)	4.0 (0.3-32.3)	4.0 (0.2-32.5)	4.0 (0.2-32.5)

Characteristics	Control (n=210)	Intervention (n=202)	Total (N=412)
IQR	2.0-8.0	2.0-8.0	2.0-8.0
Partner cohabiting, n (%)			
No	95 (45.2)	83 (41.1)	178 (43.2)
Yes	111 (52.9)	116 (57.4)	227 (55.1)
Plans for the relationship, n (%)			
End the relationship	57 (27.1)	43 (21.3)	100 (24.3)
Remain in the relationship	57 (27.1)	45 (22.3)	102 (24.8)
Unsure	96 (45.7)	114 (56.4)	210 (51.0)
Where women accessed computer, n (%)			
My house	133 (63.3)	127 (62.9)	260 (63.1)
My workplace	14 (6.7)	17 (8.4)	31 (7.5)
A family or whanau member’s house	23 (11.0)	15 (7.4)	38 (9.2)
A friend’s house	11 (5.2)	12 (5.9)	23 (5.6)
Library	13 (6.2)	13 (6.4)	26 (6.3)
Other	13 (6.2)	15 (7.4)	28 (6.8)
Violence screening criteria (among four types of violence), n (%)			
1	58 (27.6)	52 (25.7)	110 (26.7)
2-4	152 (72.4)	150 (74.3)	302 (73.3)
Thoughts of self-harm[d] , n (%)			
No	200 (95.2)	188 (93.1)	388 (94.2)
Yes	7 (3.3)	13 (6.4)	20 (4.9)
Thoughts of suicide[d] , n (%)			
No	193 (91.9)	180 (89.1)	373 (90.5)
Yes	13 (6.2)	21 (10.4)	34 (8.3)
Depression (CESD-R[e,f])			
Median (range)	31 (0-80)	34 (0-77)	32 (0-80)
IQR	18-49	17·5-51.5	18-50
Severity of violence (SVAWS[g,h])			
Median (range)	85 (47-165)	84 (46-182)	84 (46-182)
IQR	68-103	67-103.5	67.8-103.0

[a] IQR: interquartile range.

[b] Women could select one or more ethnicities.

[c] Children living in the household <18 years that woman cares for.

[d] CES-D items “wanted to hurt myself” or “wish I were dead” response “nearly every day for 2 weeks.”

[e] CESD-R: Center for Epidemiologic Studies Depression Scale, Revised.

[f] Raw scale scores without imputation. Higher scores indicate more depression symptoms; possible range 0-80).

[g] SVAWS: Severity of Violence Against Women Scale.

[h] Lower scores indicate lower exposure to violence; possible range: 46-196.

Outcomes

Intervention estimates for primary outcomes across all time periods favored the intervention (Table 2) but were not statistically or clinically significant. The SVAWS 12-month adjusted intervention estimate was −2.47 (95% CI −7.95 to 3.02) and the CESD-R 1-month adjusted intervention estimate was −0.98 (95% CI −4.89 to 2.94). No study-related adverse events were reported. Intervention estimates for secondary outcomes were also not significant (see Multimedia Appendix 2).

The statistical analysis plan included subgroup analysis of primary outcomes by ethnicity (Māori vs non-Māori) and children (responsible for one or more child vs none). There was no differential intervention effect for violence or depression symptoms based on whether women were caring for children or not.

There was a significant intervention effect for reducing violence for Māori women at 6 months (adjusted intervention estimate −14.19; 95% CI −24 to −4.37) and at 12 months (adjusted intervention estimate −12.44; 95% CI −23.35 to −1.54; Table 3 and Figure 3). There was also a significant intervention effect for reducing depression symptoms for Māori women at 3 months (adjusted intervention effect −8.7; 95% CI −15.9 to −1.6) but not at 6 or 12 months. Both violence (SVAWS) and depression (CESD-R) statistically significant changes exceeded standard error of measurement-based minimal clinically important differences [36] of 8 and 5, respectively.

Table 2. Primary outcomes.

Outcomes	Study group				Unadjusted		Adjusted[a]	
	Intervention		Control		Estimated intervention effect (95% CI)	*P* value	Estimated intervention effect (95% CI)	*P* value
	N	Mean (standard error)	N	Mean (standard error)				
Severity of violence (SVAWS[b,c])								
3 months								
Total score	163	74.92 (2.10)	168	76.60 (2.07)	−1.67 (−7.48 to 4.13)	.47	−2.69 (−7.05 to 1.66)	.23
Threat subscore	164	36.14 (1.05)	169	36.60 (1.00)	−0.46 (−3.32 to 2.39)	.64	−0.91 (−3.09 to 1.27)	.41
Acts subscore	164	29.97 (0.93)	168	31.29 (0.97)	−1.32 (−3.96 to 1.33)	.14	−1.75 (−3.77 to 0.28)	.09
Sexual aggression subscore[d]	163	8.93 (0.34)	169	8.70 (0.32)	0.23 (−0.69 to 1.14)	.53	0.95 (0.44-2.05)	.90
6 months								
Total score	162	69.36 (2.22)	180	70.88 (1.84)	−1.51 (−7.19 to 4.16)	.10	−2.16 (−7.06 to 2.73)	.39
Threat subscore	162	32.38 (1.06)	181	33.51 (0.90)	−1.13 (−3.87 to 1.61)	.15	−1.45 (−3.90 to 1.00)	.25
Acts subscore	162	28.49 (0.99)	180	29.27 (0.87)	−0.78 (−3.37 to 1.8)	.08	−0.96 (−3.18 to 1.26)	.40
Sexual aggression subscore[d]	162	8.50 (0.35)	181	8.10 (0.26)	0.40 (−0.46 to 1.25)	.95	0.69 (0.32-1.48)	.34
12 months								
Total score	173	70.00 (2.16)	183	72.43 (2.12)	−2.43 (−8.39 to 3.53)	.26	−2.47 (−7.95 to 3.02)	.38
Threat subscore	173	32.69 (1.08)	184	34.01 (1.03)	−1.32 (−4.25 to 1.62)	.22	−1.49 (−4.27 to 1.29)	.25
Acts subscore	173	28.77 (0.93)	183	30.50 (1.01)	−1.73 (−4.43 to 0.97)	.10	−1.57 (−4.01 to 0.87)	.21
Sexual aggression subscore[d]	173	8.54 (0.33)	184	8.01 (0.28)	0.53 (−0.32 to 1.38)	.44	1.09 (0.51-2.35)	.83
Depression (CESD-R[e,f])								
3 months	165	25.69 (1.52)	167	27.04 (1.56)	−1.35 (−5.64 to 2.93)	.57	−1.85 (−5.49 to 1.8)	.32
6 months	162	23.68 (1.65)	181	24.27 (1.45)	−0.59 (−4.9 to 3.73)	.53	−1.56 (−5.24 to 2.11)	.40
12 months	172	22.59 (1.63)	184	23.30 (1.51)	−0.71 (−5.08 to 3.66)	.51	−0.98 (−4.89 to 2.94)	.63

[a] All adjusted results are adjusted for baseline value of the respective score; SVAWS threat subscore was also adjusted for age group. SVAWS sexual aggression subscore and CESD-R were also adjusted for children.

[b] SVAWS: Severity of Violence Against Women Scale.

[c] Lower scores indicate lower exposure to violence.

[d] dSVAWS sexual aggression subscore was dichotomized for the adjusted analyses (low=0; high>0), and the adjusted treatment effects for this outcome are odds ratios. The full sexual aggression subscores were retained for the unadjusted analyses. CIs are based on a t-distribution using Welch approximate degrees of freedom (allowing for different variances between groups). All analyses have been carried out in the intention-to-treat analysis set. *P* values are based on Mann-Whitney test.

[e] CESD-R: Center for Epidemiologic Studies Depression Scale, Revised.

[f] Raw scale scores without imputation. Higher scores indicate more depression symptoms; possible range 0-80).

[g] SVAWS: Severity of Violence Against Women Scale.

[h] Higher scores indicate more depression symptoms.

Table 3. Primary outcomes subgroup analysis.

Outcomes per subgroup	Estimated intervention effect (95% CI)	*P* value
Severity of violence (SVAWS [a])		
Ethnicity		
3 Months		
Non-Maori	−1.09 (−6.16 to 3.98)	.67
Maori	−7.35 (−15.84 to 1.15)	.09
6 Months		
Non-Maori	1.59 (−4.02 to 7.19)	.58
Maori	−14.19 (−24 to −4.37)	.005
12 Months		
Non-Maori	0.76 (−5.57 to 7.09)	.81
Maori	−12.44 (−23.35 to −1.54)	.03
Children		
3 Months		
No Children	−2.91 (−8.8 to 2.98)	.33
Children	−2.70 (−9.24 to 3.85)	.42
6 Months		
No Children	−3.03 (−9.57 to 3.51)	.36
Children	−1.31 (−8.76 to 6.14)	.73
12 Months		
No Children	−3.1 (−10.62 to 4.4)	.42
Children	−1.65 (−9.85 to 6.55)	.69
Depression (CESD-R [b,c])		
Ethnicity		
3 Months		
Non-Maori	0.56 (−3.67 to 4.79)	.80
Maori	−8.73 (−15.88 to −1.58)	.02
6 Months		
Non-Maori	−0.51 (−4.74 to 3.72)	.81
Maori	−4.55 (−11.97 to 2.87)	.23
12 Months		
Non-Maori	1.36 (−3.16 to 5.88)	.56
Maori	−7.75 (−15.57 to 0.07)	.05
Children		
3 Months		
No Children	−0·65 (−5.55 to 4.24)	.79
Children	−1.79 (−6.7 to 3.13)	.48
6 Months		
No Children	−3.00 (−8.32 to 2.32)	.27
Children	−3.31 (−8.78 to 2.16)	.24
12 Months		
No Children	−1.25 (−6.84 to 4.33)	.66

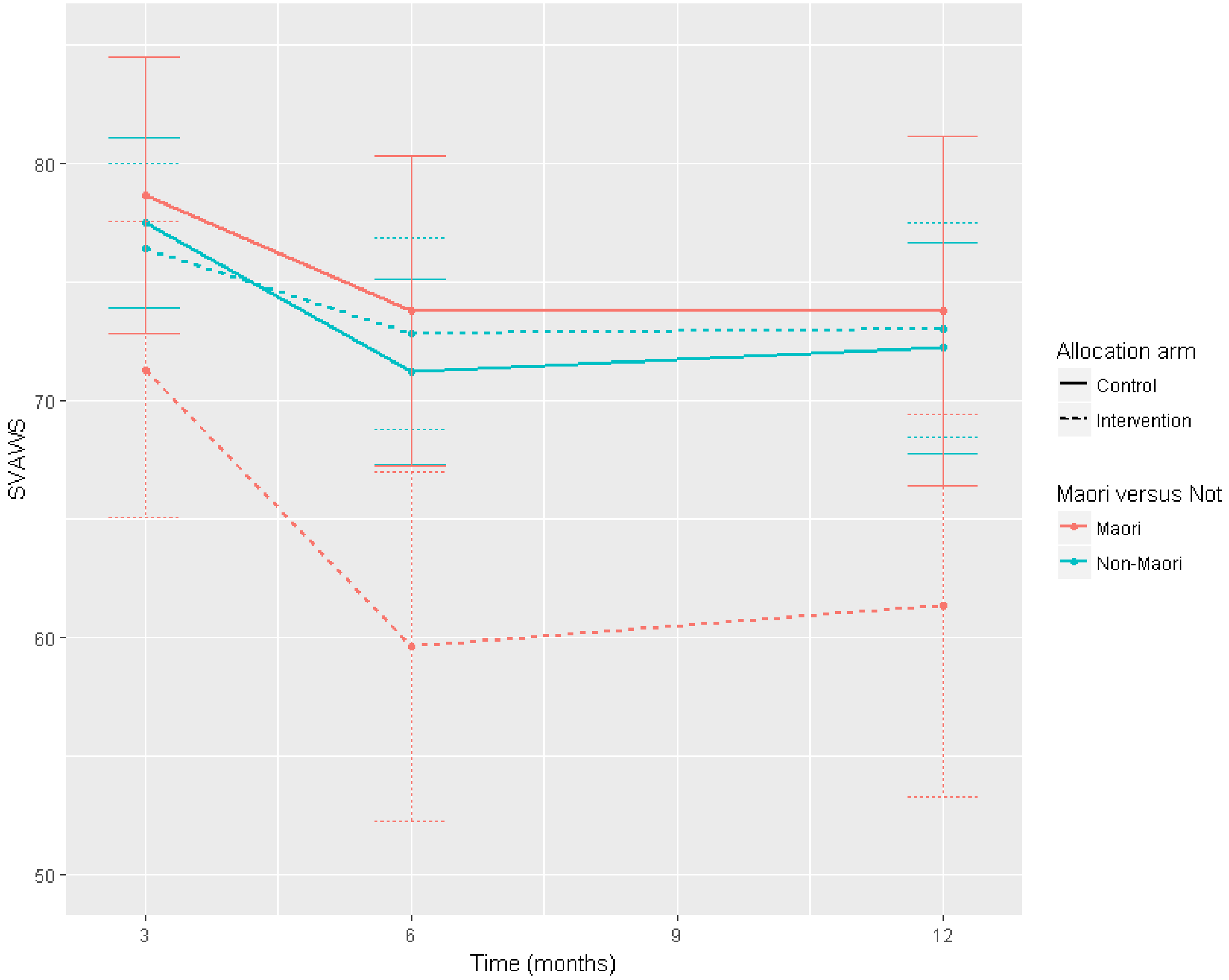
Outcomes per subgroup	Estimated intervention effect (95% CI)	*P* value
Children	1.24 (−4.55 to 7.04)	.67

[a] SVAWS: Severity of Violence Against Women Scales.

[b] CESD-R: Center for Epidemiologic Studies Depression Scale, Revised.

[c] Subgroup estimates were obtained through fitting a model including an interaction between allocation arm and Māori or children. Intervention effects are adjusted for baseline values. All analyses have been carried out in the intention-to-treat analysis set.

Figure 3. Severity of violence analysis by ethnicity and group. The Severity of Violence Against Women Scale (SVAWS) axis includes the postrandomization interquartile range values. The plotted data represent the estimated means of each group given a common starting baseline value, obtained from the fitted model.



Discussion

We tested a confidential Web-based safety decision aid (*isafe*) for women experiencing IPV, a complex health and social problem. Informed by an empowerment model [9] and decision-aid science [10], the complex intervention included three interactive components: (1) priority setting for decision making, (2) risk assessment, and (3) creation of a personalized safety action plan. In our RCT in a general New Zealand population of women 16 years of age and older, we discovered access to the regionalized *isafe* intervention was effective in reducing violence at 6 and 12 months and in reducing depression symptoms at 3 months—limited to indigenous Māori women. This is an important finding given the overrepresentation of Māori in New Zealand family violence statistics. Māori women's rate of physical and or sexual IPV in the past 12 months (14.1%) is more than three times higher than for New Zealand European women (3.9%) [37]. In addition, as in other colonized nations [38], indigenous Māori experience significant health inequities, including higher levels of unmet needs [39]. A treatment effect in a population group that experiences significant health disparities is a welcome, important finding in the struggle and moral obligation to reduce health inequities.

We offer two possible explanations for the mechanism of action for the interaction between ethnicity and outcomes. Colonization and assimilation significantly disrupted traditional Māori whānau structures and the complementary roles of men and women [40,41]. The resultant disablement of cultural practices that kept women and children safe means for many, the transmission of violence across generations has become normalized. One explanation for our findings, therefore, is that the normalization of violence in Māori whānau (families) blinds women to their risk. Completion of the *isafe* intervention brings to women's attention the level of danger they and their children are living with and empowers them to consider their priorities for safety, making them amenable to the intervention. Another explanation is that Māori women experience racism and discrimination when accessing social, health, and justice services [42]; often avoid accessing services because of the associated discrimination and other socioeconomic determinants such as cost; and are more likely, compared with non-Māori, to report unmet needs for health care [43]. As a culturally inclusive, fully automated Web-based intervention study [12-14], *isafe* avoids (but not eliminates) the potential for racism or discrimination, enabling women to access help without fear of judgment or having their children removed. Māori women may have appreciated that *isafe* was culturally inclusive rather than perpetuating the myth that IPV is only a Māori problem. This is consistent with advice from aboriginal women interviewed about addressing the problem of fetal alcohol spectrum disorders: mainstream educational strategies should be inclusive of the aboriginal community but not target them [44]. This needs to be balanced with support and advocacy for community-led Kaupapa Māori interventions by Māori for Māori [45]. This has important implications for future implementation of *isafe*.

Secondary analyses of Māori subgroup data to confirm the profile for Māori women using *isafe* (ie, does the study include Māori women at highest risk) and exploring multiple ethnicities and Māori ancestry is needed. For example, in the 2013 New Zealand census, among those who identified as Māori, half (54%) reported additional ethnicities (higher proportion for younger persons) [46]. We acknowledge that while our estimates of *isafe* impact for Māori women are significant, they are imprecise, with wide confidence intervals. The limited number of Māori participants (n=113) and essentialism in dichotomizing ethnicity in this RCT are two design constraints limiting our knowing when, for whom, and in what ways *isafe* might be beneficial to women. Alongside RCTs, research approaches such as critical realism are needed to better understand and test complex interventions for complex health and social problems such as IPV [44]. Family violence often cooccurs with a range of social and health challenges. Future work might extend the safety decision aid to address violence-related health concerns, or "join up" with other health, social, and family violence prevention initiatives.

In considering whether *isafe* is a promising intervention for a general population of women, several study limitations are important to consider. By design, all women completed baseline violence and mental health measurements (eg, SVAWS and WEB). Engaging with these measurement items may have raised women's awareness of the violence in their relationships and muted an intervention effect. In addition, outcomes were self-reported, introducing common method bias. There also may be important intermediate outcomes that we did not measure. For example, the Australian iDecide [47] and Canadian iCan [48] safety decision aid trials include general self-efficacy and self-efficacy for safety planning, respectively, in their suite of outcome measures. These two trials have not yet reported their outcomes. Finally, given the importance of considering patient-centered trial acceptability [49], women in both the intervention and control groups reported that *isafe* was useful. We had added five participant-centered acceptability questions at the 12-month follow-up midway through our trial. Among the subset of 215 women who completed their 12-month follow-up assessment after October 14, 2014, the majority agreed (true or somewhat true) the *isafe* study "provided me with new skills" (78%), "provided me with useful information" (91%), "I would suggest the site to others" (90%), "I enjoyed visiting *isafe* " (87%), and an open-ended question, "Overall, how did you find participating in the *isafe* study?" As one woman in the control group shared, *I felt protected & secure with opening up about my situation. Not judged, just helped.* Importantly, there were no differences in *isafe* acceptability by ethnicity (Māori or non-Māori). Our usefulness data suggests that although non-Māori women in the study failed to experience reduced violence or improved mental health, they did find it useful. This signals that further work is needed in understanding women's pathways to health and safety, as well as developing woman-centered measures of potential eHealth safety decision aid benefits.

This trial is one of several studies testing a Web-based safety decision aid for women experiencing IPV. Only the first trial, conducted in four US states [50], has published their findings. Among 725 consenting participants enrolled through a telephone conversation with a trained research assistant, 672 (93%) completed the 12-month assessment. Our 12-month retention rate of 87% is encouraging for future studies preferring fully automated Web-based registration. We focused our limited personal retention contacts on the 12-month assessment, particularly for 35 women who had not been automatically contacted for 3- and 6-month assessments because of a technology error. Among the three primary outcomes in the US trial, decisional conflict reduced for women in the intervention group to a greater degree than women in the control groups between the baseline pre- and baseline posttest; there were no significant differences at 6 or 12 months. The proportion of safety behaviors perceived as helpful increased 12% in the intervention group (baseline to 12 months) compared with 9% in the control group (P=.037). There were no differences between groups in the reduction of violence (or depression symptoms) over time in the US trial. Our *isafe* exploratory (not in our a priori statistical analysis plan) analysis of secondary outcomes by ethnicity suggests a larger reduction in decisional conflict among Māori women compared with non-Māori women at 3 months, but this was not statistically significant. The results are mixed in regard to safety behaviors. In particular, the Safety Behavior Checklist Helpfulness score, which resembles the measure on which a positive result was found in the US trial, shows a moderate signal favoring the control arm (on the order of 10% for Māori and 3% for non-Māori). The safety behavior activities and "what works" for women in New Zealand, particularly for Māori women, may be different compared with the United States. Further research is needed to better understand women's safety behaviors and potentially improved measurement instruments [51].

No adverse events were recorded in either the *isafe* or US trial. The *isafe* trial supports a differential effect of the Web-based safety decision aid for indigenous Māori women. As the evidence accumulates across the four current (and possibly future) international Web-based safety decision aid trials [47,48,50,52], ethical [53,54], methodological [54,55], theoretical [12,52], and practice knowledge will be gained. Importantly, these 4 studies are all examining the same intervention (with regionalization), measuring common outcomes, with follow-up of at least 12 months. This evidence will allow us to identify the magnitude of benefit and the specific populations amenable to interventions [6]. In addition to indigenous women, the safety decision aid may benefit other populations of at risk women.

Isafe was found to be a valuable early intervention resource for Māori women experiencing IPV. Dissemination of *isafe* nationwide has the potential to contribute to reducing IPV recurrence, improving mental health, and reducing injuries and perhaps femicide amongst Māori women exposed to IPV. In addition, *isafe* is likely to be a cost effective, economically sustainable resource. In 2014, family violence was conservatively estimated to cost New Zealand NZ $4.1 billion per annum, with NZ $2,198 in costs avoided for every New Zealander whose experience of family violence is prevented [56]. Cost analyses are needed to calculate the cost per woman using the decision aid against the yearly cost of maintaining advocacy and counseling services. In developing a sustainable open access online platform for *isafe*, ongoing modification to keep pace with technology advancements, people’s use of technology, and future evidence will be needed. *isafe* may also be a useful resource to include in formal health [4], police, and justice IPV responses to women experiencing violence, alongside the necessary services working with abusive partners [57]. Dissemination studies using the Reach, Effectiveness, Adoption, Implementation, and Maintenance framework [58] for example, are needed to guide evidence-based practice across these intervention locations.

In conclusion, the interactive Web-based *isafe* decision aid includes risk assessment, priority setting for decision making, and creation of a personalized safety action plan. Our findings provide impetus for further dissemination and testing of interactive, individualized Web-based interventions to reduce IPV and associated health harms in at risk populations.

Acknowledgments

The authors are appreciative of the many women who were willing to engage in the *isafe* study, to community partners who provided pathways to safety for study participants, and members of the DMC. They are thankful for Rangi Davis, psychosynthesis practitioner, MNZAC, Te Haehae Healing Services, for her ongoing guardianship of the *isafe* study regarding Ao Māori. They acknowledge James Case, the senior software engineer who developed and managed the technology for the *isafe* study. All authors take responsibility for the accuracy and completeness of the reported data and vouch for the fidelity of the trial to the protocol. This work was funded by the New Zealand Health Research Council. The funder had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Authors' Contributions

JKM, AV, DW, SNR, NG, and KE participated in the conception, study design, and obtaining of funding. TD and CM contributed to the data collection and data processing. AV and RS contributed to data analysis. JKM, AV, DW, and SNR contributed to drafting of the manuscript. JKM, AV, DW, SNR, NG, KE, TD, CM, and RS contributed to the interpretation of the results and made critical revisions to the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Outcome assessment schedule for the *isafe* trial.

[PDF File (Adobe PDF File), 23KB - jmir_v19i12e426_app1.pdf]

Multimedia Appendix 2

Secondary outcomes.

[PDF File (Adobe PDF File), 35KB - jmir_v19i12e426_app2.pdf]

Multimedia Appendix 3

CONSORT‐EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 11MB - jmir_v19i12e426_app3.pdf]

References

1. Campbell JC. Health consequences of intimate partner violence. Lancet 2002 Apr 13;359(9314):1331-1336. [doi: 10.1016/S0140-6736(02)08336-8] [Medline: 11965295]
2. Garcia-Moreno C, DeVries K, Stockl H, Pallitto C, Watts C, Abrahams N. Global and regional estimates of violence against women: Prevalence and health effects of intimate partner violence and non-partner sexual violence. Geneva: World Health Organization; 2013.
3. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, 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]
4. García-Moreno C, Hegarty K, d'Oliveira AF, Koziol-McLain J, Colombini M, Feder G. The health-systems response to violence against women. The Lancet 2015 Apr;385(9977):1567-1579. [doi: 10.1016/S0140-6736(14)61837-7]
5. World Health Organization. Global plan of action to strengthen the role of the health system within a national multisectoral response to address interpersonal violence, in particular against women and girls, and against children. Geneva: World Health Organization; 2016.
6. Rivas C, Ramsay J, Sadowski L, Davidson LL, Dunne D, Eldridge S, et al. Advocacy interventions to reduce or eliminate violence and promote the physical and psychosocial well-being of women who experience intimate partner abuse. Cochrane Database Syst Rev 2015(12):CD005043. [doi: 10.1002/14651858.CD005043.pub3] [Medline: 26632986]
7. Fanslow JL, Robinson EM. Help-seeking behaviors and reasons for help seeking reported by a representative sample of women victims of intimate partner violence in New Zealand. J Interpers Violence 2010 May;25(5):929-951. [doi: 10.1177/0886260509336963] [Medline: 19597160]
8. Glass N, Eden KB, Bloom T, Perrin N. Computerized aid improves safety decision process for survivors of intimate partner violence. J Interpers Violence 2009 Dec 29;25(11):1947-1964. [doi: 10.1177/0886260509354508]
9. Dutton MA. Empowering and healing the battered woman: a model for assessment and intervention. New York: Springer Publishing Company; 1992.
10. Stacey D, Légaré F, Col NF, Bennett CL, Barry MJ, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2014;1:CD001431. [doi: 10.1002/14651858.CD001431.pub4] [Medline: 24470076]
11. Eden KB, Perrin NA, Hanson GC, Messing JT, Bloom TL, Campbell JC, et al. Use of online safety decision aid by abused women. Am J Prev Med 2015 Apr;48(4):372-383. [doi: 10.1016/j.amepre.2014.09.027]
12. Young-Hauser AM, Eden KB, Wilson D, Koziol-Mclain J. Intimate partner violence: modifying an internet-based safety decision aid to a New Zealand context. J Technol Hum Serv 2014 Nov 26;32(4):297-311. [doi: 10.1080/15228835.2014.967905]
13. Koziol-McLain J, Vandal AC, Nada-Raja S, Wilson D, Glass NE, Eden KB, et al. A web-based intervention for abused women: the New Zealand isafe randomised controlled trial protocol. BMC Public Health 2015;15:56 [FREE Full text] [doi: 10.1186/s12889-015-1395-0] [Medline: 25637195]
14. Koziol-McLain J, McLean C, Rohan M, Sisk R, Dobbs T, Nada-Raja S, et al. Participant recruitment and engagement in automated eHealth trial registration: challenges and opportunities for recruiting women who experience violence. J Med Internet Res 2016 Oct 25;18(10):e281 [FREE Full text] [doi: 10.2196/jmir.6515] [Medline: 27780796]
15. Saaty T. Decision making with the analytic hierarchy process. IJSSCI 2008;1(1):83-98. [doi: 10.1504/IJSSci.2008.017590]
16. Campbell JC, Webster DW, Glass N. The danger assessment: validation of a lethality risk assessment instrument for intimate partner femicide. J Interpers Violence 2009 Apr;24(4):653-674. [doi: 10.1177/0886260508317180] [Medline: 18667689]
17. Glass N, Perrin N, Hanson G, Bloom T, Gardner E, Campbell JC. Risk for reassault in abusive female same-sex relationships. Am J Public Health 2008 Jun;98(6):1021-1027. [doi: 10.2105/ajph.2007.117770]
18. Eaton WW, Smith C, Ybarra M, Muntaner C, Tien A. Center for Epidemiologic Studies Depression Scale: review and revision (CESD and CESD-R). In: Maruish ME, editor. The Use Of Psychological Testing For Treatment Planning And Outcomes third Edition, Volume 3. Mahwah: Lawrence Erlbaum Associates; 2004.
19. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. Appl Psychol Meas 2016 Jul 26;1(3):385-401. [doi: 10.1177/014662167700100306]
20. Van Dam NT, Earleywine M. Validation of the Center for Epidemiologic Studies Depression Scale--Revised (CESD-R): pragmatic depression assessment in the general population. Psychiatry Res 2011 Mar 30;186(1):128-132. [doi: 10.1016/j.psychres.2010.08.018] [Medline: 20843557]
21. Marshall L. Development of the severity of violence against women scales. J Fam Viol 1992 Jun;7(2):103-121. [doi: 10.1007/BF00978700]
22. Smith PH, Earp JA, DeVellis R. Measuring battering: development of the Women's Experience with Battering (WEB) Scale. Womens Health 1995;1(4):273-288. [Medline: 9373384]
23. Blanchard EB, Jones-Alexander J, Buckley TC, Forneris CA. Psychometric properties of the PTSD Checklist (PCL). Behav Res Ther 1996 Aug;34(8):669-673. [Medline: 8870294]

24. Norris FH, Hamblen JL. Standardized self-report measures of civilian trauma and PTSD. In: Wilson JP, Keane TM, editors. Assessing Psychological Trauma and PTSD, 2nd ed. New York: The Guilford Press; 2004:63-102.
25. Reinert DF, Allen JP. The alcohol use disorders identification test: an update of research findings. Alcohol Clin Exp Res 2007 Feb;31(2):185-199. [doi: 10.1111/j.1530-0277.2006.00295.x] [Medline: 17250609]
26. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption--II. Addiction 1993 Jun;88(6):791-804. [Medline: 8329970]
27. Smith PC, Schmidt SM, Allensworth-Davies D, Saitz R. A single-question screening test for drug use in primary care. Arch Intern Med 2010 Jul 12;170(13):1155-1160 [FREE Full text] [doi: 10.1001/archinternmed.2010.140] [Medline: 20625025]
28. Yudko E, Lozhkina O, Fouts A. A comprehensive review of the psychometric properties of the Drug Abuse Screening Test. J Subst Abuse Treat 2007 Mar;32(2):189-198. [doi: 10.1016/j.jsat.2006.08.002] [Medline: 17306727]
29. O'Connor AM. Validation of a decisional conflict scale. Med Decis Making 1995;15(1):25-30. [Medline: 7898294]
30. Sullivan CM, Bybee DI. Reducing violence using community-based advocacy for women with abusive partners. J Consult Clin Psychol 1999 Feb;67(1):43-53. [Medline: 10028208]
31. Parker B, McFarlane J, Soeken K, Silva C, Reel S. Testing an intervention to prevent further abuse to pregnant women. Res Nurs Health 1999 Feb;22(1):59-66. [Medline: 9928964]
32. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. Br Med J 2011 Feb 07;342:d40 [FREE Full text] [Medline: 21300711]
33. White IR, Thompson SG. Adjusting for partially missing baseline measurements in randomized trials. Stat Med 2005 Apr 15;24(7):993-1007. [doi: 10.1002/sim.1981] [Medline: 15570623]
34. Carpenter J, Kenward M. Missing data in randomised controlled trials - a practical guide. London: London School of Hygiene & Tropical Medicine; 2007.
35. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. J Stat Soft 2015;67(1):-. [doi: 10.18637/jss.v067.i01]
36. Wyrwich KW, Nienaber NA, Tierney WM, Wolinsky FD. Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. Med Care 1999 May;37(5):469-478. [Medline: 10335749]
37. Fanslow J, Robinson E, Crengle S, Perese L. Juxtaposing beliefs and reality: prevalence rates of intimate partner violence and attitudes to violence and gender roles reported by New Zealand women. Violence Against Women 2010 Jul;16(7):812-831. [doi: 10.1177/1077801210373710] [Medline: 20558771]
38. Kirmayer LJ, Brass G. Addressing global health disparities among Indigenous peoples. Lancet 2016 Jul 09;388(10040):105-106. [doi: 10.1016/S0140-6736(16)30194-5] [Medline: 27108233]
39. Ministry of Health. The health of Māori adults and children. Wellington: New Zealand Ministry of Health; 2013.
40. Kruger T, Pitman M, Grennell D, McDonald T, Mariu D, Pomare A. Transforming Whanau Violence: A Conceptual Framework (2nd ed). Wellington: Te Puni Kokiri (Ministry of Maori Affairs); 2004.
41. Dobbs T, Eruera M. Kaupapa Māori wellbeing framework: the basis for whānau violence prevention and intervention. Auckland: New Zealand Family Violence Clearinghouse; 2014.
42. Harris R, Tobias M, Jeffreys M, Waldegrave K, Karlsen S, Nazroo J. Effects of self-reported racial discrimination and deprivation on Māori health and inequalities in New Zealand: cross-sectional study. Lancet 2006 Jun 17;367(9527):2005-2009. [doi: 10.1016/S0140-6736(06)68890-9] [Medline: 16782491]
43. Ministry of Health. Annual update of key results 2015/2016: New Zealand Health Survey. Wellington: Ministry of Health; 2016.
44. Blagg H, Bluett-Boyd N, Williams E. Innovative models in addressing violence against Indigenous women: state of knowledge paper. Sydney: ANROWS; 2015.
45. Smith L. Decolonizing methodologies: research and indigenous peoples. London: Zed Books; 1999.
46. Statistics New Zealand. Archive.stats. 2013. 2013 Census ethnic group profiles: Māori URL: http://archive.stats.govt.nz/Census/2013-census/profile-and-summary-reports/ethnic-profiles.aspx?request_value=24705&tabname=Age,sex,andethnicities
47. Hegarty K, Tarzia L, Murray E, Valpied J, Humphreys C, Taft A, et al. Protocol for a randomised controlled trial of a web-based healthy relationship tool and safety decision aid for women experiencing domestic violence (I-DECIDE). BMC Public Health 2015 Aug 01;15:736 [FREE Full text] [doi: 10.1186/s12889-015-2072-z] [Medline: 26231225]
48. Ford-Gilboe M, Varcoe C, Scott-Storey K, Wuest J, Case J, Currie LM, et al. A tailored online safety and health intervention for women experiencing intimate partner violence: the iCAN Plan 4 Safety randomized controlled trial protocol. BMC Public Health 2017 Mar 21;17(1):273 [FREE Full text] [doi: 10.1186/s12889-017-4143-9] [Medline: 28327116]
49. Kost RG, Lee LM, Yessis J, Wesley RA, Henderson DK, Coller BS. Assessing participant-centered outcomes to improve clinical research. N Engl J Med 2013 Dec 05;369(23):2179-2181 [FREE Full text] [doi: 10.1056/NEJMp1311461] [Medline: 24304050]
50. Glass NE, Perrin NA, Hanson GC, Bloom TL, Messing JT, Clough AS, et al. The longitudinal impact of an internet safety decision aid for abused women. Am J Prev Med 2017 May;52(5):606-615. [doi: 10.1016/j.amepre.2016.12.014] [Medline: 28108189]

51. Wilson D, Jackson D, Herd R. Confidence and connectedness: Indigenous Māori women's views on personal safety in the context of intimate partner violence. Health Care Women Int 2016 Jul;37(7):707-720. [doi: 10.1080/07399332.2015.1107069] [Medline: 26491900]
52. Tarzia L, Murray E, Humphreys C, Glass N, Taft A, Valpied J, et al. I-DECIDE: an online intervention drawing on the Psychosocial Readiness Model for women experiencing domestic violence. Womens Health Issues 2016;26(2):208-216. [doi: 10.1016/j.whi.2015.07.011] [Medline: 26362841]
53. Allenby R, Dobbs T, Diesfeld K, Nada Raja S, Wilson D, Koziol-McLain J. Safety in online research with women experiencing intimate partner violence: what about the children? Ethics Behav 2016 Mar 30;27(1):1-17. [doi: 10.1080/10508422.2016.1140579]
54. Tarzia L, Valpied J, Koziol-McLain J, Glass N, Hegarty K. Methodological and ethical challenges in a web-based randomized controlled trial of a domestic violence intervention. J Med Internet Res 2017 Mar 28;19(3):e94 [FREE Full text] [doi: 10.2196/jmir.7039] [Medline: 28351830]
55. Ford-Gilboe M, Wathen CN, Varcoe C, MacMillan HL, Scott-Storey K, Mantler T, et al. Development of a brief measure of intimate partner violence experiences: the Composite Abuse Scale (Revised)-Short Form (CASR-SF). BMJ Open 2016 Dec 07;6(12):e012824 [FREE Full text] [doi: 10.1136/bmjopen-2016-012824] [Medline: 27927659]
56. Kahui S, Snively S. Measuring the Economic Costs of Child Abuse and Intimate Partner Violence to New Zealand. Wellington: MoreMedia Enterprises; 2014.
57. Family Violence Death Review Committee. Fifth Report Data: January 2009 to December 2015. Wellington: Health Quality & Safety Commission; 2017.
58. Gaglio B, Shoup JA, Glasgow RE. The RE-AIM framework: a systematic review of use over time. Am J Public Health 2013 Jun;103(6):e38-e46. [doi: 10.2105/AJPH.2013.301299] [Medline: 23597377]

Abbreviations

CESD-R: Center for Epidemiologic Studies Depression Scale-Revised
DA: Danger Assessment
DA-R: Danger Assessment-Revised
DMC: data monitoring committee
eHealth: electronic health
IPV: intimate partner violence
ITT: intention-to-treat
IQR: interquartile range
RCT: randomized controlled trial
SVAWS: Severity of Violence Against Women Scale
WEB: Women’s Experience with Battering

Edited by G Eysenbach; submitted 01.08.17; peer-reviewed by K Clinical Practice - Assessment, L Tarzia, A Hattery; comments to author 24.08.17; revised version received 15.11.17; accepted 30.11.17; published 10.01.18.

Please cite as:
Koziol-McLain J, Vandal AC, Wilson D, Nada-Raja S, Dobbs T, McLean C, Sisk R, Eden KB, Glass NE
Efficacy of a Web-Based Safety Decision Aid for Women Experiencing Intimate Partner Violence: Randomized Controlled Trial
J Med Internet Res 2018;19(12):e426
URL: http://www.jmir.org/2017/12/e426/
doi:10.2196/jmir.8617
PMID:29321125

Original Paper

Testimonials and Informational Videos on Branded Prescription Drug Websites: Experimental Study to Assess Influence on Consumer Knowledge and Perceptions

Helen W Sullivan[1], PhD, MPH; Amie C O'Donoghue[1], PhD; Jennifer Gard Read[2], MPH; Jacqueline B Amoozegar[2], MSPH; Kathryn J Aikin[1], PhD; Douglas J Rupert[2], MPH

[1] US Food and Drug Administration, Silver Spring, MD, United States
[2] RTI International, Research Triangle Park, NC, United States

Corresponding Author:
Helen W Sullivan, PhD, MPH
US Food and Drug Administration
10903 New Hampshire Ave, Bldg 51
Silver Spring, MD, 20993
United States
Phone: 1 301 796 1200
Fax: 1 301 847 8444
Email: helen.sullivan@fda.hhs.gov

Abstract

Background: Direct-to-consumer (DTC) promotion of prescription drugs can affect consumer behaviors and health outcomes, and Internet drug promotion is growing rapidly. Branded drug websites often capitalize on the multimedia capabilities of the Internet by using videos to emphasize drug benefits and characteristics. However, it is unknown how such videos affect consumer processing of drug information.

Objective: This study aimed to examine how videos on prescription drug websites, and the inclusion of risk information in those videos, influence consumer knowledge and perceptions.

Methods: We conducted an experimental study in which online panel participants with acid reflux (n=1070) or high blood pressure (n=1055) were randomly assigned to view 1 of the 10 fictitious prescription drug websites and complete a short questionnaire. On each website, we manipulated the type of video (patient testimonial, mechanism of action animation, or none) and whether the video mentioned drug risks.

Results: Participants who viewed any video were less likely to recognize drug risks presented only in the website text (*P*≤.01). Including risk information in videos increased participants’ recognition of the risks presented in the videos (*P*≤.01). However, in some cases, including risk information in videos decreased participants’ recognition of the risks *not* presented in the videos (ie, risks presented in text only; *P*≤.04). Participants who viewed a video without drug risk information thought that the website placed more emphasis on benefits, compared with participants who viewed the video with drug risk information (*P*≤.01). Compared with participants who viewed a video without drug risk information, participants who viewed a video with drug risk information thought that the drug was less effective in the high blood pressure sample (*P*=.03) and thought that risks were more serious in the acid reflux sample (*P*=.01). There were no significant differences between risk and nonrisk video conditions on other perception measures (*P*>.05). In addition, we noted a few differences among the types of videos.

Conclusions: Including risks in branded drug website videos may increase in-video risk retention at the expense of text-only risk retention.

(J Med Internet Res 2018;20(1):e13) doi:10.2196/jmir.7959

KEYWORDS

Internet; marketing; drug prescriptions; risk

Introduction

Background

Direct-to-consumer (DTC) marketing of prescription drugs is correlated with increases in consumer drug spending, prescription drug use, and prescription requests [1-6]. However, studies have revealed mixed findings as to whether DTC marketing leads to more informed decisions by consumers [7-9]. DTC marketing on the Internet is rapidly growing [10,11] with a 648% increase in spending on Internet promotion between 2001 and 2010 [12]; therefore, it is critical to understand how this medium may affect outcomes related to informed decision making.

Pharmaceutical companies use a number of different Internet activities, most commonly to promote products and communicate with consumers [13]. To this end, most leading pharmaceutical companies maintain two types of websites: websites with corporate business information and websites promoting specific drugs or medications, known as branded drug websites (eg, DrugX.com). Branded drug websites typically provide information about the branded drug, the disease or medical condition, support measures, and ways to locate physicians and pharmacists [14]. Branded drug websites often use videos as a marketing tool to emphasize a promoted drug's benefits and characteristics [14-18]. These videos may share features with DTC television advertising. For instance, DTC television advertisements frequently include factual information and positive emotional appeals, often in the form of testimonials [19]. However, consumers may react differently to videos in Internet promotion [20]. One way in which videos on websites differ from television advertisements is that these videos are often presented along with text on the website, creating a mix of dynamic and static formats. The movement in the video could increase attention to the video, thereby causing greater recall of the information in the video and perhaps influencing the perceptions of the risks or the benefits. Previous research on the influence of dynamic videos is mixed [21], but a meta-analysis [22] found a small benefit for dynamic versus static images when learning information.

Testimonial Videos

Branded drug websites have sometimes showcased video testimonials of expert medical sources as well as patients who have been treated successfully with the promoted drug [15,16]. Previous research has not examined testimonials within the context of DTC marketing; however, studies in other health-related areas have found mixed evidence about how patient testimonials affect individuals' perceptions of their disease risk. Some studies have found that testimonials heighten consumer risk perceptions for issues such as human papillomavirus (HPV), human immunodeficiency virus (HIV), and other sexually transmitted diseases [23,24]. However, other studies have found that testimonials did not change consumer risk perceptions on these issues [25]. Trevana and colleagues reviewed 3 articles that used narratives to convey risk information and concluded that patient testimonials should be used with caution because they can have unpredictable effects on risk perceptions [26]. In addition, another study concluded that the use of patient testimonials in decision aids should be avoided until their impact is better understood, because there are concerns that stories have the potential to bias patients' decisions [27]. Given these mixed findings, we examined the role of video testimonials on branded drug websites, where the understanding of drug risks and benefits is a critical part of an informed discussion with a health care provider.

Informational Videos

Another strategy to promote branded drugs on the Internet is the use of videos that provide information about the product. Although previous studies have examined the presentation of information about topics ranging from astrophysics [28] to pulley systems [21], we are unaware of any study that has examined the presentation of prescription drug information on branded drug websites. DTC promotion is replete with technical information that must be conveyed to adequately represent the benefits and risks of the product, according to US Food and Drug Administration (FDA) regulations [29]. Thus, we examined the presence of a dynamic informational video on the website describing the drug's mechanism of action to determine whether this presentation influenced viewers' understanding of the promoted drug's benefits and risks.

Research Questions

FDA regulations state that prescription drug promotion should include a fair balance of information about the benefits and risks of promoted products, in terms of both content and presentation [29]. The regulations further specify that important risk information should be presented in at least the audio, or in both the audio and visual of broadcast advertisements. All prescription drug promotions should be truthful, balanced, and nonmisleading, regardless of the media in which that promotion occurs [30]. However, questions remain about how best to achieve "fair balance" in Internet DTC promotion [31-34]. The purpose of this study was to examine how videos on prescription drug websites—and the content of those videos—influence consumers' knowledge, perceptions, and intentions related to the advertised drug.

We had 2 overarching research questions: (1) how does the presence of Web videos (testimonials and informational videos) influence consumer knowledge and perceptions of drugs? and (2) how does including risk information in Web videos influence consumer knowledge and perceptions of drugs? We hypothesized that the presence, versus absence, of Web videos would lead to greater retention of benefit information, and therefore higher drug efficacy perceptions and greater intentions to search for information about the drug. This hypothesis is based on the expectation that video presentations draw attention and that the focus of videos is often benefit information (eg, centered on a patient attesting to the drug effectiveness). Accordingly, we also hypothesized that the presence, versus absence, of Web videos would lead to less retention of the risk information, and therefore lower drug risk perceptions. In addition, we hypothesized that videos including risk information, compared with videos without risk information, would lead to greater retention of the risk information, and therefore higher drug risk perceptions and lower intentions to search for information about the drug. We did not have hypotheses about potential differences between testimonials and informational videos; however, because we randomly assigned participants to the type of video they viewed, we tested for any differences between these conditions. Finally, we explored whether including a video and whether including risk information in the video influenced perceptions of the website itself.

Methods

Study Design

We conducted the study with consumers with acid reflux and consumers with high blood pressure and exposed them to a mock drug website advertising a fictitious drug (Fentiva or Plistaz) designed to treat their respective condition. We chose these medical conditions because they affect a large number of people and represent both symptomatic (acid reflux) and asymptomatic (high blood pressure) conditions. We manipulated the type of video participants viewed and the prominence of the risk information on the websites. All websites contained the benefit and risk information in the text. We manipulated whether the websites presented a video of a personal testimonial, an informational mechanism of action video, or no video at all (control group). In the video conditions, we also manipulated the prominence of the risk information by including some risk information (high-risk prominence) or by not including some risk information (low-risk prominence) in the video. Experimental conditions are presented in Figure 1. To ensure that our fictitious websites were realistic, we reviewed actual prescription drug websites and consulted FDA’s Office of Prescription Drug Promotion. We designed the fictitious websites to mirror real-life prescription drug websites in structure (eg, homepage plus 2 subpages), content (eg, drug benefits and risks and tips for disease management), and design (eg, consistent banner and heading across all pages and photos of patients or caregivers). The study was approved by the relevant institutional review boards.

Participants

The sampling frame was the GfK Custom Research North America KnowledgePanel, a probability-based online consumer panel based on a representative random sample of US adults. Panelists were randomly selected and invited to participate in the study if they were aged 18 years or older, self-reported to the panel that they were medically diagnosed with acid reflux or high blood pressure, did not participate in the study pretest, were capable of viewing and listening to websites on a desktop computer or device, had broadband Internet access, and used a device with Adobe Flash Player software. We randomly assigned participants to 1 of the 5 experimental arms within each illness group and confirmed their self-reported medical diagnosis. For acid reflux, we invited a total of 2226 panelists to participate in the study. Out of these panelists, 1774 responded and 1070 completed the study. For high blood pressure, we invited a total of 2020 panelists to participate in the study. Out of these panelists, 1559 responded and 1055 completed the study. A summary of participants’ demographic characteristics is presented in Table 1.

Procedure

After completing the screening questions, participants were instructed to click a hyperlink to open the study website in a new window (675×1064 pixels). In addition, participants were instructed to turn up the volume and watch any videos. Exposure to the website was forced (ie, participants were not able to proceed to the questionnaire without clicking the hyperlink and landing on the website). Each website included a homepage and 2 subpages. Participants were allowed to navigate back and forth among them without time restrictions. Internet browser controls were removed from the stimuli window to simplify navigation.

In all conditions, the website presented both the drug’s risk and benefit information as text. In the control condition, no video was present on the website. In the testimonial condition, the website included a video featuring an actor depicting a patient who described how the drug worked for them (including treating acid reflux or lowering blood pressure). In the mechanism of action video condition, the website included a dynamic animated video depicting how the drug mechanism treats the illness (including treating acid reflux or lowering blood pressure). In the high-risk prominence testimonial condition, the “patient” stated some of the risk information. In the high-risk prominence mechanism of action videos, the voiceover stated some of the risk information while the risks were presented dynamically through text or icons. The risk information included in the high-risk prominence conditions was identical in the testimonials and mechanism of action videos. The website layout is depicted in Figure 2.

Figure 1. Experimental conditions.

Table 1. Demographic characteristics of the participants.

Demographic characteristics	Acid reflux sample, weighted (N=1070) n (%)	High blood pressure sample, weighted (N=1055) n (%)
Sex		
Male	449 (41.96)	505 (47.87)
Female	621 (58.04)	550 (52.13)
Age, years		
18-24	22 (2.06)	6 (0.57)
25-34	69 (6.45)	26 (2.46)
35-44	145 (13.55)	94 (8.91)
45-54	212 (19.81)	188 (17.82)
55-64	274 (25.61)	313 (29.67)
65-74	249 (23.27)	310 (29.38)
75+	99 (9.25)	118 (11.18)
Race or ethnicity [a]		
White	925 (86.45)	838 (79.43)
Black	97 (9.06)	194 (18.39)
Other	105 (9.81)	80 (7.58)
Hispanic	88 (8.22)	93 (8.82)
Education		
Less than high school	52 (4.86)	50 (4.74)
High school graduate	419 (39.16)	408 (38.67)
Some college	315 (29.44)	305 (28.91)
College degree or higher	284 (26.54)	292 (27.68)

[a] There is some overlap between categories as Hispanics are also counted in the 3 race categories.

We programmed the videos to play automatically, thus forcing exposure to the manipulations. No control bar was available; participants could not stop or mute the video once it began playing, although they could close the website whenever they wanted. Participants could replay the video as many times as desired. Once participants closed the stimuli window and continued to the questionnaire, they were unable to view the website again.

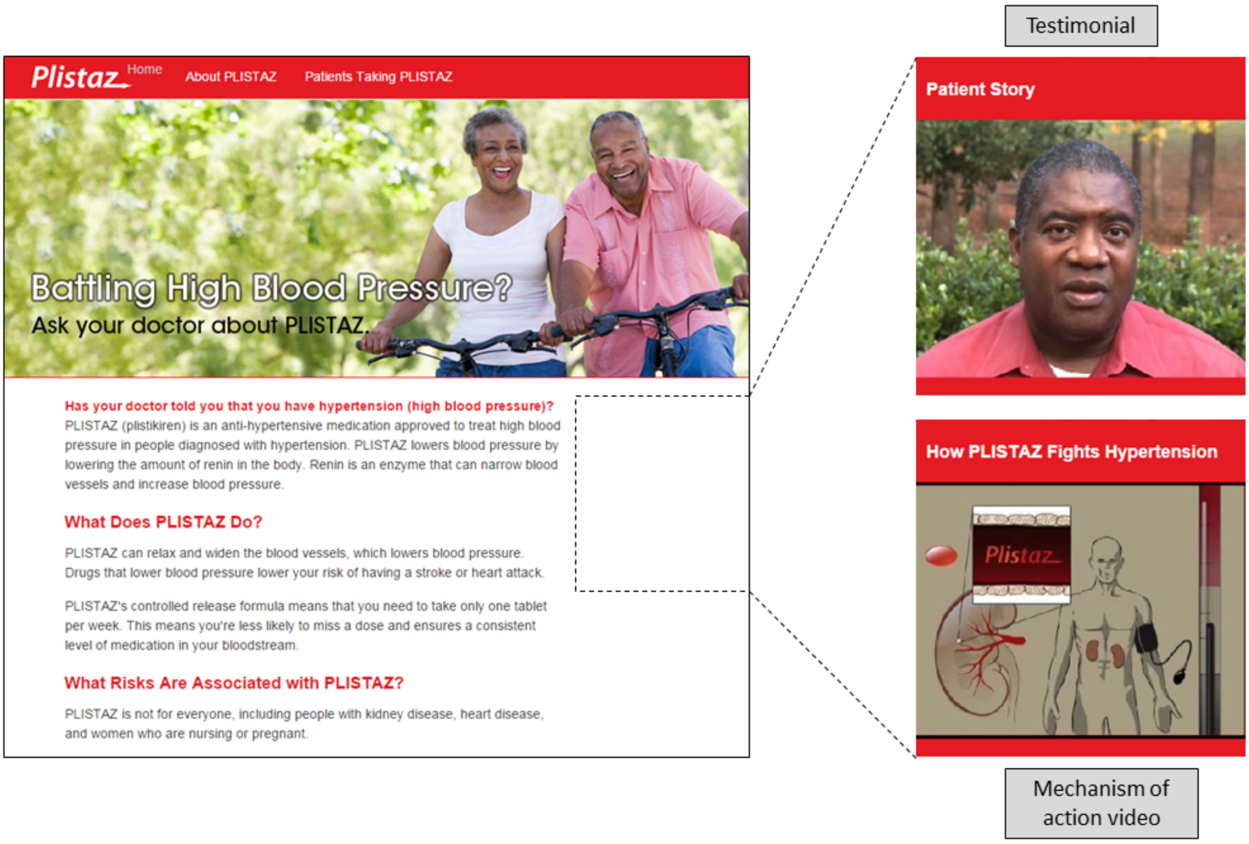
Measures

We used Web logs to track participants' interaction with the fictitious drug websites and translated these logs into navigation variables that could be used to analyze participants' behavior, including two measures of video exposure: (1) whether participants were fully exposed or partially exposed to the video and (2) whether or not they replayed the video.

We measured risk recall by asking participants to list the risks of the drug in an open-ended text box. For the acid reflux sample, we created a measure of risk recall by coding whether participants reported the risks presented in the video: nausea, headache, stomach pain, diarrhea, constipation, or fractures (0-6). For the high blood pressure sample, we created a measure of risk recall by coding whether participants reported the side effects (diarrhea, rash, or cough) and limitations (that the drug could not be taken if pregnant, used as a salt substitute, used with salt, or used while drinking alcohol) presented in the video (0-7).

To measure risk recognition, we presented participants with two risks that appeared in the video and in the website text, two risks that appeared in the website text only, and four filler statements. Participants indicated whether each statement was mentioned on the website as a risk of taking the drug. For the acid reflux sample, we measured whether they recognized the risks presented in the video [fracture risk (0-1) and nausea risk (0-1)] and the risks presented in the website text only [warnings and precautions regarding special liver tests (0-1) and women who are nursing (0-1)]. Note that although the videos did not include the warning about special liver tests, they did say to tell your doctor if you have liver disease. For the high blood pressure sample, we measured whether they recognized the risks presented in the video [diarrhea risk (0-1) and salt-intake warning (0-1)] and the risks presented in the website text only [fetal risk (0-1) and angioedema warning (0-1)].

Figure 2. Example of study stimuli.



Has your doctor told you that you have hypertension (high blood pressure)?
PLISTAZ (plistikiren) is an anti-hypertensive medication approved to treat high blood pressure in people diagnosed with hypertension. PLISTAZ lowers blood pressure by lowering the amount of renin in the body. Renin is an enzyme that can narrow blood vessels and increase blood pressure.

What Does PLISTAZ Do?

PLISTAZ can relax and widen the blood vessels, which lowers blood pressure. Drugs that lower blood pressure lower your risk of having a stroke or heart attack.

PLISTAZ's controlled release formula means that you need to take only one tablet per week. This means you're less likely to miss a dose and ensures a consistent level of medication in your bloodstream.

What Risks Are Associated with PLISTAZ?

PLISTAZ is not for everyone, including people with kidney disease, heart disease, and women who are nursing or pregnant.

We measured benefit recall by asking participants to list the benefits of the drug in an open-ended text box. For the acid reflux sample, we created a measure of benefit recall by coding whether participants reported that the drug could relieve acid reflux [including terms such as heartburn and GERD (gastroesophageal reflux disease); (0-1)]. For the high blood pressure sample, we created a measure of benefit recall by coding whether participants reported that the drug could lower blood pressure (0-1).

To measure benefit recognition, we presented participants with a list of statements about the drug. Participants indicated whether each statement was mentioned on the website as a benefit of taking the drug. For the acid reflux sample, we measured whether the participants recognized the benefit of the drug, “Fentiva can provide relief from heartburn” (0-1). For the high blood pressure sample, we measured whether the participants recognized the benefit of the drug, “Plistaz lowers blood pressure by reducing the amount of renin in the body” (0-1).

Perceived drug risk was measured by 2 items that assessed participants’ thoughts on how many people would have side effects out of 100 people who take the drug (*likelihood*; open-ended item with values ranging from 0 to 100) and how serious the side effects would be for them (*magnitude*; 1=not at all serious, 6 = very serious). Perceived drug efficacy was measured by 2 items that assessed participants’ thoughts on how many people would benefit out of 100 people who take the drug (*likelihood*; open-ended item with values ranging from 0 to 100) and how effective the drug would be for them (*magnitude*; 1=would help very little to 6=would help a lot).

Perceived balance of drug benefits and risks was measured by asking participants how they would rate the drug on its balance of risks and benefits (1=risks outweigh benefits to 7=benefits outweigh risks).

We measured two types of behavioral intentions: the intention to interact with one’s physician and the intention to seek additional information about the drug on the Internet. For the physician intention measure, we averaged 3 items that assessed how likely participants were to talk to their doctor about the drug, ask their doctor for a sample of the drug, and ask their doctor to prescribe the drug (acid reflux alpha=.94; high blood pressure alpha=.93). For the information-seeking intention measure, we averaged 3 items that assessed how likely participants were to look for information about the drug on medical websites, look for information on the Internet about people’s experience with the drug, and print information from the drug website (acid reflux alpha=.90; high blood pressure alpha=.90). All intention items used the same scale (1=very unlikely to 5=very likely).

We also asked participants about the website itself. We measured website skepticism with the average of 2 items that assessed whether participants thought the information on the websites was true (acid reflux alpha=.66; high blood pressure alpha=.71). Perceived balance of website benefits and risks was measured by asking participants whether they thought the website placed more emphasis on risks or benefits of the drug (1=more emphasis on risks to 7=more emphasis on benefits).

Data Analysis

Weighting was used to account for the underrepresentation of minority groups and other types of sampling and survey errors. We transformed *perceived drug efficacy likelihood* (by squaring it) and *perceived drug risk likelihood* (by using the natural log), resulting in approximately normal distribution of the data in the two illness populations. We conducted all hypothesis tests using the transformed variables.

We conducted one-way analysis of variances (ANOVAs) to test the relationship between video type condition (testimonial, mechanism of action, and control) and the risk recall, perception, intention, and skepticism measures. If the effect was significant at P<.05, we performed pairwise comparisons by comparing the testimonial and mechanism of action video type conditions with the control condition, using a Bonferroni-adjusted threshold of P=.025. We conducted two-way ANOVAs to test the relationship between the video type condition (testimonial and mechanism of action) and risk prominence (low and high) and the risk recall, perception, intention, and skepticism measures. We conducted two sets of logistic regressions for risk recognition and benefit recall and retention: one examining the video type conditions compared with the control condition and one examining the main effects and interaction of video type condition and risk prominence.

Results

Video Exposure

Acid Reflux and High Blood Pressure

All participants were exposed to the video, and most of them viewed the entire video (94.8% in the acid reflux sample and 98.6% in the high blood pressure sample). Only a few participants replayed the video (7.5% in the acid reflux sample and 6.8% in the high blood pressure sample).

Retention of Drug Risk Information

Acid Reflux

We tested the effect of risk prominence and video type condition on risk recall. We found that participants in the high-risk prominence condition recalled more risks compared with participants in the low-risk prominence condition ($F_{1,860}$=4.00, P=.046, d=.14; Table 2).

We also tested the effect of risk prominence and video type condition on the risk recognition measures. We found a significant effect of risk prominence on two of the risk recognition measures, and a significant interaction on a third measure. Compared with participants in the low-risk prominence condition, participants in the high-risk prominence condition were more likely to recognize the risk of fracture presented in the video and the risk of special liver tests that was alluded to in the video but presented only in the website text ($F_{1,860}$=6.54, P=.01, d=.17 and $F_{1,860}$=7.88, P=.01, d=.19, respectively). There was a significant interaction for the nursing warning presented only in the website text ($F_{1,860}$=10.38, P=.001, d=.22). In the mechanism of action conditions, 83.8% (standard error [SE] 3.0) of participants in the high-risk prominence condition and 75.2% (SE 4.7) in the low-risk prominence condition recognized the nursing warning; however, this was reversed in the testimonial conditions (high-risk prominence=64.1% [SE 5.2] and low-risk prominence=83.1% [SE 3.6]).

Table 2. Weighted percentages and means (standard errors) by risk prominence.

Risk retention variables	AR[a] sample		HBP[b] sample	
	Low-risk prominence (n=426)	High-risk prominence (n=435)	Low-risk prominence (n=419)	High-risk prominence (n=404)
Risk recall[c], mean (SE)	1.80 (0.12)[d]	2.13 (0.11)	1.02 (0.09)	1.12 (0.08)
Risk recognition: video, % (SE)				
AR fracture risk and HBP diarrhea risk	77.5 (3.4)[d]	88.4 (2.5)	66.6 (3.7)[d]	86.9 (2.7)
AR nausea risk and HBP salt intake warning	77.5 (3.2)	77.2 (3.1)	68.0 (3.3)	71.1 (3.6)
Risk recognition: text, % (SE)				
AR liver warning and HBP fetal risk	65.4 (3.7)[d]	79.3 (3.2)	69.3 (3.5)[d]	58.7 (3.7)
AR nursing warning and HBP angioedema warning	79.1 (3.0)	73.9 (3.1)	63.0 (3.9)[d]	49.0 (3.7)
Benefit recall, % (SE)	44.2 (3.9)	38.3 (3.5)	70.1 (3.6)	62.3 (3.7)
Benefit recognition, % (SE)	85.6 (2.7)	82.7 (2.8)	88.4 (2.7)	80.3 (3.3)

[a] AR: acid reflux.

[b] HBP: high blood pressure.

[c] Risk recall: 0-6 correct in the AR sample and 0-7 in the HBP sample.

[d] Significantly different from the high-risk prominence condition, P<.05.

On comparing the video type conditions with the control condition on risk recall and risk recognition, we found effects for one risk recognition measure. Compared with participants in the mechanism of action and testimonial conditions, participants in the control condition were more likely to recognize the nursing warning presented only in the website text ($F_{1,1069}$=6.05, P=.01, d=.15 and $F_{1,1069}$=12.18, P=.001, d=.21, respectively).

The risk prominence and video type conditions did not significantly differ on the risk recognition measure regarding the nausea risk presented in the video (P>.05).

High Blood Pressure

We tested the effect of risk prominence and video type condition on risk recall. We found that participants in the mechanism of action condition recalled more risks compared with participants in the testimonial condition ($F_{1,822}$=5.04, P=.03, d=.16; Table 3).

We also tested the effect of risk prominence and video type condition on the risk recognition measures. We found a significant effect of risk prominence on two of the risk recognition measures. Compared with participants in the low-risk prominence condition, participants in the high-risk prominence condition were more likely to recognize the risk of diarrhea presented in the video ($F_{1,822}$=17.22, P<.001, d=.29). However, they were *less* likely to recognize the fetal and angioedema risks presented only in the website text ($F_{1,822}$=4.30, P=.04, d=.14 and $F_{1,822}$=6.39, P=.01, d=.18, respectively).

On comparing the video type conditions with the control condition on risk recall and risk recognition, we found effects for one risk recognition measure. Participants in the control condition were more likely to recognize the angioedema risk presented only in the website text compared with participants in the testimonial condition ($F_{1,1054}$=12.33, P<.001, d=.22).

The risk prominence and video type condition did not significantly differ on the risk recognition measure regarding the salt intake risk presented in the video (P>.05).

Retention of Drug Benefit Information

Acid Reflux

We tested the effect of risk prominence and video type condition on benefit recall and benefit recognition. We found no significant effects for benefit recall (P>.05). We found significant effects for benefit recognition when we compared the video type conditions. Specifically, participants in the testimonial and control conditions were more likely to recognize the drug's benefit compared with participants in the mechanism of action conditions ($F_{1,860}$=15.22, P<.001, d=.26 and $F_{1,1069}$=11.74, P=.001, d=.21, respectively).

High Blood Pressure

We tested the effect of risk prominence and video type conditions on benefit recall and benefit recognition. We found significant effects for benefit recall and benefit recognition when we compared the video type conditions. Participants in the mechanism of action condition were more likely to recall the drug's benefit compared with participants in the testimonial and control conditions ($F_{1,822}$=5.27, P=.02, d=.16 and $F_{1,1054}$=7.50, P=.01, d=.17, respectively). They were also more likely to recognize the drug's benefit compared with participants in the testimonial condition ($F_{1,822}$=6.01, P=.01, d=.17).

Table 3. Weighted percentages and means (standard errors) by video type condition.

Risk retention variables	AR[a] sample			HBP[b] sample		
	Mechanism of action (n=430)	Testimonial (n=431)	Control (n=209)	Mechanism of action (n=411)	Testimonial (n=412)	Control (n=232)
Risk recall[c], mean (SE)	1.90 (0.11)	2.03 (0.12)	1.91 (0.15)	1.20[d] (0.09)	0.94 (0.08)	1.13 (0.11)
Risk recognition: video, % (SE)						
AR fracture risk and HBP diarrhea risk	83.7 (3.0)	82.4 (3.1)	88.8 (3.7)	79.1 (3.0)	74.0 (3.6)	80.6 (4.0)
AR nausea risk and HBP salt intake warning	75.0 (3.4)	79.6 (2.9)	80.1 (4.5)	68.9 (3.3)	70.1 (3.6)	74.4 (4.4)
Risk recognition: text, % (SE)						
AR liver warning and HBP fetal risk	68.6 (3.7)	76.2 (3.1)	66.2 (5.3)	63.8 (3.5)	35.6 (3.8)	75.9 (4.3)
AR nursing warning and HBP angioedema warning	79.5 (2.8)[e]	73.5 (3.3)[e]	89.9 (2.6)	61.2 (3.7)	51.1 (3.9)[e]	74.3 (4.7)
Benefit recall, % (SE)	41.7 (3.6)	40.7 (3.8)	55.1 (5.3)	72.1 (3.3)[d,e]	60.5 (3.9)	55.3 (5.3)
Benefit recognition, % (SE)	77.4 (3.1)[d,e]	90.9 (2.1)	93.1 (2.3)	89.8[d] (2.3)	79.1 (3.6)	81.0 (4.0)

[a] AR: acid reflux.

[b] HBP: high blood pressure.

[c] Risk recall: 0-6 correct in the AR sample and 0-7 in the HBP sample.

[d] Significantly different from the testimonial condition, P<.05.

[d] Significantly different from the control condition; Bonferroni-adjusted for two comparisons with the control condition, P<.025.

Table 4. Weighted means (standard errors) by risk prominence.

Dependent variables	Acid reflux sample, mean (SE)		High blood pressure sample, mean (SE)	
	Low-risk prominence	High-risk prominence	Low-risk prominence	High-risk prominence
Perceived drug risk				
Likelihood[a]	33.11 (1.92)	31.93 (1.69)	33.34 (1.83)	37.29 (1.89)
Magnitude[b]	3.46[c] (0.09)	3.76 (0.08)	3.75 (0.09)	3.73 (0.08)
Perceived drug efficacy				
Likelihood[a]	70.66 (1.50)	71.43 (1.47)	64.48 (1.80)	65.25 (1.69)
Magnitude[d]	4.84 (0.08)	4.95 (0.07)	4.79[c] (0.08)	4.52 (0.09)
Perceived balance of drug benefits and risks[e]	4.51 (0.11)	4.42 (0.11)	4.49 (0.08)	4.33 (0.10)
Intention[f]				
Physician interaction	2.88 (0.09)	2.74 (0.09)	2.55 (0.10)	2.29 (0.10)
Search on the Internet	2.50 (0.09)	2.39 (0.08)	2.38 (0.10)	2.12 (0.09)
Website skepticism[g]	3.49 (0.09)	3.45 (0.08)	3.77 (0.08)	3.75 (0.10)
Perceived balance of website benefit and risk information[h]	5.16[c] (0.10)	4.72 (0.10)	5.14[c] (0.10)	4.78 (0.10)

[a] Perceived drug risk and efficacy likelihood: 0-100 people. Although transformations of perceived drug risk and efficacy likelihood were used in analyses, the untransformed weighted means are presented here for the ease of interpretation.

[b] Perceived drug risk magnitude: 1 (not at all serious) to 6 (very serious).

[c] Significantly different from the high-risk prominence condition, $P<.05$.

[d] Perceived drug efficacy magnitude: 1 (help a little) to 6 (help a lot).

[e] Perceived balance of drug benefits and risks: 1 (risks outweigh benefits) to 7 (benefits outweigh risks).

[f] Intention: 1 (very unlikely) to (5=very likely).

[g] Website skepticism: 1 (extremely unlikely) to 7 (extremely likely).

[h] Perceived balance of website benefit and risk information: 1 (more emphasis on risks) to 7 (more emphasis on benefits).

Perceived Drug Efficacy

Acid Reflux

We tested the effect of risk prominence and video type condition on the perceived drug efficacy measures. We found no significant effects for perceived drug efficacy likelihood or magnitude ($P>.05$; Tables 4 and 5).

High Blood Pressure

We tested the effect of risk prominence and video type condition on the perceived drug efficacy measures. We found a significant effect of risk prominence on one of the measures. Specifically, participants in the low-risk prominence condition thought that the drug would work better compared with participants in the high- risk prominence condition ($F_{1,810}=4.60$, $P=.03$, $d=.15$). All other effects for the perceived drug efficacy measures were nonsignificant ($P>.05$).

Perceived Drug Risk

Acid Reflux

We tested the effect of risk prominence and video type condition on perceived drug likelihood and magnitude. For perceived drug magnitude, we found a significant effect when comparing video type conditions with the control condition and a significant effect of risk prominence. Participants in the control condition thought that the drug's side effects and negative outcomes would be more serious compared with participants in the mechanism of action condition ($F_{1,1048}=7.47$, $P=.01$, $d=.17$). In addition, participants in the high-risk prominence condition thought that the drug's side effects and negative outcomes would be more serious compared with participants in the low-risk prominence condition ($F_{1,841}=6.40$, $P=.01$, $d=.17$). We found no significant effects for perceived drug risk likelihood ($P>.05$).

High Blood Pressure

We tested the effect of risk prominence and video type condition on perceived drug likelihood and magnitude. We found no significant effects ($P>.05$).

Perceived Balance of Drug Benefits and Risks

Acid Reflux

We tested the effect of risk prominence and video type condition on the perceived balance of drug benefits and risks. We found no significant effects ($P>.05$).

Table 5. Weighted means (standard errors) by video type condition.

Dependent variables	Acid reflux sample, mean (SE)			High blood pressure sample, mean (SE)		
	Mechanism of action	Testimonial	Control	Mechanism of action	Testimonial	Control
Perceived drug risk						
Likelihood[a]	34.40 (1.95)	30.67 (1.62)	36.08 (2.69)	37.25 (1.90)	33.21 (1.78)	35.98 (2.49)
Magnitude[b]	3.52[c] (0.08)	3.70 (0.10)	3.93 (0.13)	3.72 (0.09)	3.76 (0.09)	3.89 (0.12)
Perceived drug efficacy						
Likelihood[a]	72.38 (1.34)	69.70 (1.61)	74.97 (2.09)	66.48 (1.40)	63.20 (2.02)	64.23 (1.98)
Magnitude[d]	4.88 (0.07)	4.92 (0.08)	5.02 (0.10)	4.71 (0.08)	4.60 (0.09)	4.50 (0.09)
Perceived balance of drug benefits and risks[e]	4.50 (0.10)	4.43 (0.12)	4.62 (0.15)	4.57[f] (0.08)	4.25 (0.10)	4.37 (0.12)
Intention[g]						
Physician interaction	2.81 (0.08)	2.81 (0.10)	2.83 (0.13)	2.48 (0.10)	2.37 (0.10)	2.12 (0.12)
Search on the Internet	2.39 (0.08)	2.48 (0.09)	2.50 (0.12)	2.36 (0.10)	2.14 (0.09)	2.11 (0.11)
Website skepticism[h]	3.35[f] (0.07)	3.58[c] (0.09)	3.24 (0.12)	3.70 (0.09)	3.82 (0.09)	3.69 (0.11)
Perceived balance of website benefit and risk information[i]	5.05 (0.09)	4.83 (0.11)	4.69 (0.16)	5.07[c] (0.10)	4.85[c] (0.10)	4.27 (0.16)

[a] Perceived drug risk and efficacy likelihood: 0-100 people. Although transformations of perceived drug risk and efficacy likelihood were used in analyses, the untransformed weighted means are presented here for ease of interpretation.

[b] Perceived drug risk magnitude: 1 (not at all serious) to 6 (very serious).

[c] Significantly different from the control condition; Bonferroni-adjusted for two comparisons with the control condition, $P<.025$.

[d] Perceived drug efficacy magnitude: 1 (help a little) to 6 (help a lot).

[e] Perceived balance of drug benefits and risks: 1 (risks outweigh benefits) to 7 (benefits outweigh risks).

[f] Significantly different from the testimonial condition, $P<.05$.

[g] Intentions: 1 (very unlikely) to 5 (very likely).

[h] Website skepticism: 1 (extremely unlikely) to 7 (extremely likely).

[i] Perceived balance of website benefit and risk information: 1 (more emphasis on risks) to 7 (more emphasis on benefits).

High Blood Pressure

We tested the effect of risk prominence and video type condition on the perceived balance of drug benefits and risks. We found a significant effect of video type condition. Participants in the mechanism of action condition thought that the benefits outweighed the risks compared with participants in the testimonial condition ($F_{1,809}=6.25$, $P=.01$, $d=.18$).

Behavioral Intentions

Acid Reflux and High Blood Pressure

We tested the effect of risk prominence and video type condition on physician interaction and Internet search intentions. In both the acid reflux and high blood pressure samples, we found no significant effects ($P>.05$).

Website Skepticism

Acid Reflux

We tested the effect of risk prominence and video type condition on website skepticism. We found a significant effect of video type condition and a significant interaction. Participants in the testimonial condition were more skeptical of the website compared with participants in the control and mechanism of action conditions ($F_{1,1063}=5.47$, $P=.02$, $d=.14$ and $F_{1,856}=4.14$, $P=.04$, $d=.14$, respectively). A significant interaction with risk prominence ($F_{1,856}=4.63$, $P=.03$) suggests that this was driven by the low-risk prominence conditions (testimonial, low-prominence condition mean=3.72, SE 0.12; mechanism of action video, low-prominence condition mean=3.25, SE 0.11; testimonial, high-prominence condition mean=3.44, SE 0.13; mechanism of action video, high-prominence condition mean=3.46, SE 0.10).

High Blood Pressure

We tested the effect of risk prominence and video type conditions on website skepticism. We found no significant effects ($P>.05$).

Perceived Balance of Website Benefit and Risk Information

Acid Reflux

We tested the effect of risk prominence and video type condition on the perceived balance of website benefit and risk information. The effect of risk prominence was significant; participants in the low-risk prominence condition thought that the website placed more emphasis on benefits compared with participants in the high-risk prominence condition ($F_{1,857}$=9.50, P=.002, d=.21).

High Blood Pressure

We tested the effect of risk prominence and video type condition on the perceived balance of website benefit and risk information. We found significant effects when comparing video type conditions with the control condition, and a significant effect of risk prominence. Participants in the mechanism of action and testimonial conditions thought that the website placed more emphasis on benefits compared with participants in the control condition ($F_{1,1037}$=17.56, P<.001, d=.26 and $F_{1,1037}$=9.35, P=.002, d=.19, respectively). Participants in the low-risk prominence condition thought that the website placed more emphasis on benefits compared with participants in the high-risk prominence condition ($F_{1,809}$=6.17, P=.01, d=.17).

Discussion

Principal Findings

This study investigated how videos embedded on branded prescription drug websites influence consumers' knowledge, perceptions, and intentions. We tested websites with no video, with a testimonial, and with an informational video describing the drug's mechanism of action. Within the testimonial and mechanism of action videos, we manipulated the prominence of risk information, with some risk information included, or not included, in the videos (with risk information always present in the text of the website). The results of this study suggest that embedded videos can affect consumers' knowledge of drug information but are less likely to affect perceptions and intentions.

We found that including risk information in videos increased participants' recognition of the risks presented in the videos. However, in some cases, including risk information in videos decreased participants' recognition of the risks *not* presented in the videos (ie, risks presented in text only). The Communication-Human Information Processing Model states that for risk information to be understood, individuals must first switch their attention to the information and then maintain their attention on the information [35,36]. The videos may capture and maintain their attention; therefore, when videos are presented on websites, the inclusion of risks in the videos can increase the retention of some of those risks. At the same time, it may decrease the retention of risks not mentioned in the video by preventing individuals from switching and maintaining their attention on the text. Participants may assume that all risks are included in the video, and therefore, they may not read the risk information provided only in the website's text.

Including risks in the video did not affect participants' retention of the drug's main benefit, although in the high blood pressure sample, it did decrease the perceived magnitude of drug efficacy. In addition, we found that not including risk information in the videos shifted participants' views of the website, causing them to believe that the website placed more emphasis on benefits. In most cases, however, this did not translate into increased skepticism of the website; only participants in the acid reflux sample who saw the testimonial without risk information reported more skepticism. These findings suggest that if videos are used, the inclusion of risk information in videos can increase consumers' knowledge of the risks while not diminishing consumers' knowledge of the drug's benefit or making them more skeptical.

We found that the videos had no effect on intentions and a limited effect on risk perceptions; in the acid reflux sample only, including risk information in the videos increased the perceived magnitude of the drug's risks. Quantifying or personalizing the risks in testimonials may be necessary to change perceptions. These findings are also consistent with previous research demonstrating that consumers consider numerous sources of information about prescription drugs—particularly, health care providers—and that their perceptions of and intentions about specific medications are influenced by multiple factors, including their personal health history, satisfaction with current treatment, and knowledge of others' experiences with the drugs [37-43]. Consequently, Web videos promoting a prescription drug may not be powerful or persuasive enough to shift individuals' perceptions and intentions unless they are part of a larger marketing campaign or align with a health care provider's recommendation to take a medication.

The study samples consisted of individuals who had been diagnosed with one of two health conditions—acid reflux and high blood pressure. Across medical conditions, we saw a similar pattern of effects for risk prominence, such that risk prominence affected risk retention but had little or no effect on benefit retention and perceptions. This provides some confidence that these results would generalize across different medical conditions and different website and video executions. However, the effects of video type condition were not consistent across medical conditions. The two video styles examined in this study—testimonials and mechanism of action videos—were not entirely equivalent, and the executions differed across medical conditions. To be realistic and to mirror actual content on branded drug websites, the videos contained different benefit information. The mechanism of action videos focused on how each fictitious drug worked; the testimonials presented an individual patient's experience benefiting from the fictitious drug. This distinction means that differences between testimonial and mechanism of action video type conditions in the study's findings could be caused by multiple factors, such as video style (eg, live action vs animated), video content, or even video duration. The benefit retention results reflect this; in the acid reflux sample, the mechanism of action video decreased the retention of the drug's benefits, whereas it increased retention in the high blood pressure sample. Future research should standardize content within testimonials and mechanism of action videos to determine whether differences are attributable to visuals or content. This would provide a less realistic, but more controlled, setting in which these concepts can be tested.

Limitations

This study was a controlled experiment with realistic stimuli, large sample sizes, and high statistical power. Nevertheless, the study has several limitations that should be considered when interpreting the findings. First, the study samples were limited to individuals with household broadband Internet access and, thus, may contain a higher proportion of white, older, and more educated individuals than the US adult population. The study was also limited to participants with two illness conditions and did not include individuals without health issues. One avenue for future research is to test these concepts in a different sample. For instance, future research could focus on individuals actively seeking prescription treatments (eg, newly diagnosed and dissatisfied with current treatment). Future research should also examine whether key differences between these two illness conditions (eg, symptomatic vs asymptomatic, many nonprescription treatment options vs few nonprescription treatment options) explain why certain effects were present in only one population.

Individuals may be passively exposed to television and print DTC advertising, whereas individuals who visit DTC websites may be actively seeking information about treatments. Thus, participants in our study may have had less interest in and paid less attention to the study DTC website compared with individuals actively seeking information on the Internet. Conducting future research with individuals actively seeking prescription treatments would partially address this limitation as well.

Conclusions

Our results reflect the caution in using testimonials urged by Elwyn and colleagues [27]. Including a video on a prescription drug website can affect website credibility (for instance, by increasing skepticism and changing the perceived balance of information on the website). It can also enhance or detract from consumers' knowledge of the drug's benefits and risks. When videos are used on prescription drug websites, the inclusion of risk information in the videos can lead to greater knowledge of the product's important risk information. Thus, "fair balance" may be enhanced by including risk information in website videos.

Acknowledgments

This study was funded by a contract from the US Food and Drug Administration, Office of Prescription Drug Promotion.

The authors would like to acknowledge Kayla Gray, Scott Boggs, Elizabeth Robbins, Maria Ashbaugh, Cliff Haac, and John Bollenbacher (RTI International) for their help with developing the stimuli and Rebecca Moultrie (RTI International) and Bridget Kelly (RTI International) for their assistance in creating questionnaires and collecting data. The authors would also like to thank Sarah Kandefer, Grier Page, Annette Green, and Paul Mosquin (RTI International) for their help with the analysis.

Conflicts of Interest

None declared.

References

1. Avery RJ, Eisenberg MD, Simon KI. The impact of direct-to-consumer television and magazine advertising on antidepressant use. J Health Econ 2012 Sep;31(5):705-718. [doi: 10.1016/j.jhealeco.2012.05.002] [Medline: 22835472]
2. Daubresse M, Hutfless S, Kim Y, Kornfield R, Qato DM, Huang J, et al. Effect of direct-to-consumer advertising on asthma medication sales and healthcare use. Am J Respir Crit Care Med 2015 Jul 01;192(1):40-46 [FREE Full text] [doi: 10.1164/rccm.201409-1585OC] [Medline: 25879303]
3. Government Accountability Office. Prescription drugs: Improvements needed in FDA's oversight of direct-to-consumer advertising. Washington, DC: US Government Accountability Office; 2006.
4. Kim Y, Kornfield R, Shi Y, Vera L, Daubresse M, Alexander GC, et al. Effects of televised direct-to-consumer advertising for varenicline on prescription dispensing in the United States, 2006-2009. Nicotine Tob Res 2016 May;18(5):1180-1187. [doi: 10.1093/ntr/ntv198] [Medline: 26385926]
5. Layton JB, Kim Y, Alexander GC, Emery SL. Association between direct-to-consumer advertising and testosterone testing and initiation in the United States, 2009-2013. JAMA 2017 Dec 21;317(11):1159-1166 [FREE Full text] [doi: 10.1001/jama.2016.21041] [Medline: 28324090]
6. Niederdeppe J, Byrne S, Avery RJ, Cantor J. Direct-to-consumer television advertising exposure, diagnosis with high cholesterol, and statin use. J Gen Intern Med 2013 Jul;28(7):886-893 [FREE Full text] [doi: 10.1007/s11606-013-2379-3] [Medline: 23463454]
7. Frosch D, Grande D. Direct-to-consumer advertising of prescription drugs. LDI Issue Brief 2010;15(3):1-4. [Medline: 20469541]
8. Grenard JL, Uy V, Pagán JA, Frosch DL. Seniors' perceptions of prescription drug advertisements: a pilot study of the potential impact on informed decision making. Patient Educ Couns 2011 Oct;85(1):79-84 [FREE Full text] [doi: 10.1016/j.pec.2010.10.008] [Medline: 21044826]
9. Mintzes B. Advertising of prescription-only medicines to the public: does evidence of benefit counterbalance harm? Annu Rev Public Health 2012 Apr;33:259-277. [doi: 10.1146/annurev-publhealth-031811-124540] [Medline: 22429162]
10. Mackey TK, Cuomo RE, Liang BA. The rise of digital direct-to-consumer advertising?: comparison of direct-to-consumer advertising expenditure trends from publicly available data sources and global policy implications. BMC Health Serv Res 2015 Jun 19;15:236 [FREE Full text] [doi: 10.1186/s12913-015-0885-1] [Medline: 26084705]

11. Sullivan HW, Aikin KJ, Chung-Davies E, Wade M. Prescription drug promotion from 2001-2014: data from the U.S. food and drug administration. PLoS One 2016;11(5):e0155035 [FREE Full text] [doi: 10.1371/journal.pone.0155035] [Medline: 27149513]
12. Kornfield R, Donohue J, Berndt ER, Alexander GC. Promotion of prescription drugs to consumers and providers, 2001-2010. PLoS One 2013;8(3):e55504 [FREE Full text] [doi: 10.1371/journal.pone.0055504] [Medline: 23469165]
13. Dupuits FM. The effects of the internet on pharmaceutical consumers and providers. Dis Manag Health Out 2002;10(11):679-691. [doi: 10.2165/00115677-200210110-00002]
14. Sheehan KB. In poor health: an assessment of privacy policies at direct-to-consumer web sites. J Public Policy Mark 2005 Sep;24(2):273-283. [doi: 10.1509/jppm.2005.24.2.273]
15. Griffiths KM, Christensen H, Evans K. Pharmaceutical company websites as sources of information for consumers. Dis Manag Health Out 2002;10(4):205-214. [doi: 10.2165/00115677-200210040-00001]
16. Ledford CJ. Content analysis of internet marketing strategies: how pharmaceutical companies communicate about contraceptives with consumers online. Soc Mar Q 2009 Jul 15;15(sup1):55-71. [doi: 10.1080/15245000903038308]
17. Macias W, Lewis LS. A content analysis of direct-to-consumer (DTC) prescription drug web sites. J Advertising 2003 Dec;32(4):43-56. [doi: 10.1080/00913367.2003.10639147]
18. Pashupati K. Beavers, bubbles, bees, and moths. JAR 2009 Sep 01;49(3):373-393. [doi: 10.2501/S0021849909090473]
19. Frosch DL, Krueger PM, Hornik RC, Cronholm PF, Barg FK. Creating demand for prescription drugs: a content analysis of television direct-to-consumer advertising. Ann Fam Med 2007;5(1):6-13 [FREE Full text] [doi: 10.1370/afm.611] [Medline: 17261859]
20. Logan K. And now a word from our sponsor: do consumers perceive advertising on traditional television and online streaming video differently? J Mark Commun 2012;19(4):258-276 [FREE Full text] [doi: 10.1080/13527266.2011.631568]
21. Boucheix J, Schneider E. Static and animated presentations in learning dynamic mechanical systems. Learn Instr 2009 Apr;19(2):112-127. [doi: 10.1016/j.learninstruc.2008.03.004]
22. Höffler TN, Leutner D. Instructional animation versus static pictures: a meta-analysis. Learn Instr 2007 Dec;17(6):722-738. [doi: 10.1016/j.learninstruc.2007.09.013]
23. Mevissen FEF, Meertens RM, Ruiter RAC, Feenstra H, Schaalma HP. HIV/STI risk communication: the effects of scenario-based risk information and frequency-based risk information on perceived susceptibility to chlamydia and HIV. J Health Psychol 2009 Jan;14(1):78-87. [doi: 10.1177/1359105308097948] [Medline: 19129340]
24. Nan X, Dahlstrom MF, Richards A, Rangarajan S. Influence of evidence type and narrative type on HPV risk perception and intention to obtain the HPV vaccine. Health Commun 2015;30(3):301-308. [doi: 10.1080/10410236.2014.888629] [Medline: 25061715]
25. Pot M, van Keulen HM, Paulussen TWGM, Otten W, Ruiter RA. Statistical versus narrative risk information: implications for future communication about HPV vaccination. European Health Psychologist 2014;16:579.
26. Trevena LJ, Zikmund-Fisher BJ, Edwards A, Gaissmaier W, Galesic M, Han PK, et al. Presenting quantitative information about decision outcomes: a risk communication primer for patient decision aid developers. BMC Med Inform Decis Mak 2013;13 Suppl 2:S7 [FREE Full text] [doi: 10.1186/1472-6947-13-S2-S7] [Medline: 24625237]
27. Elwyn G, O'Connor A, Stacey D, Volk R, Edwards A, Coulter A, et al. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. BMJ 2006 Aug 26;333(7565):417 [FREE Full text] [doi: 10.1136/bmj.38926.629329.AE] [Medline: 16908462]
28. Lewalter D. Cognitive strategies for learning from static and dynamic visuals. Learn Instr 2003 Apr;13(2):177-189. [doi: 10.1016/S0959-4752(02)00019-1]
29. US Food and Drug Administration. 2016. Prescription-drug advertisements 21 CFR 202.1 URL: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?fr=202.1 [WebCite Cache ID 6vpZDT1pX]
30. U.S. Food and Drug Administration. FDA. 2014. Guidance for Industry: internet/social media platforms with character space limitations—presenting risk and benefit information for prescription drugs and medical devices URL: https://www.fda.gov/downloads/drugs/guidances/ucm401087.pdf [accessed 2017-12-28] [WebCite Cache ID 6w2olfLBg]
31. Hoy MG, Park JS. Principles in action: an examination of food and drug administration letters involving violative internet promotions from 1997 to 2012. J Public Policy Mark 2014 Nov;33(2):127-142. [doi: 10.1509/jppm.12.097]
32. Mintzes B. The tip of the iceberg of misleading online advertising comment on “trouble spots in online direct-to-consumer prescription drug promotion: a content analysis of FDA warning letters”. Int J Health Policy Manag 2016 Feb 18;5(5):329-331. [doi: 10.15171/ijhpm.2016.19]
33. Southwell BG, Rupert DJ. future challenges and opportunities in online prescription drug promotion research comment on “trouble spots in online direct-to-consumer prescription drug promotion: a content analysis of FDA warning letters”. Int J Health Policy Manag 2016 Jan;5(3):211-213 [FREE Full text] [doi: 10.15171/ijhpm.2016.05] [Medline: 26927597]
34. U.S. Food and Drug Administration. FDA. 2014. Public hearing on promotion of FDA-regulated medical products using the internet and social media tools URL: https://www.fda.gov/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/ucm184250.htm [accessed 2017-12-19] [WebCite Cache ID 6vpYOLpbI]
35. Conzola V, Wogalter M. A communication-human information processing (C-HIP) approach to warning effectiveness in the workplace. J Risk Res 2001;4(4):309-322. [doi: 10.1080/13669870110062712]

36. Wogalter M. Communication-human information processing (C-HIP) model. In: Handbook of Warnings. Mahwah, NJ: Lawrence Erlbaum Associates; 2006:51-61.
37. Chen C, Wu J, Yen M, Chen Z. A model of medication-taking behavior in elderly individuals with chronic disease. J Cardiovasc Nurs 2007;22(5):359-365. [doi: 10.1097/01.JCN.0000287029.87746.58] [Medline: 17724416]
38. DeLorme DE, Huh J, Reid LN. Seniors' perceptions of prescription drug information sources. Int J Pharm Healthc Mark 2007 Jul 03;1(2):107-127. [doi: 10.1108/17506120710762979]
39. Doucette WR, Schommer JC. Consumer preferences for drug information after direct-to-consumer advertising. Drug Information Journal 1998 Jan 01;32(4):1081-1088. [doi: 10.1177/009286159803200428]
40. Gregory R, Peters E, Slovic P. Making decisions about prescription drugs: a study of doctor–patient communication. Health Risk Soc 2011 Jun;13(4):347-371. [doi: 10.1080/13698575.2011.575455]
41. Kim WJ, King KW. Product category effects on external search for prescription and nonprescription drugs. J Advert 2009 Apr;38(1):5-20. [doi: 10.2753/JOA0091-3367380101]
42. Lau DT, Briesacher BA, Mercaldo ND, Halpern L, Osterberg EC, Jarzebowski M, et al. Older patients' perceptions of medication importance and worth: an exploratory pilot study. Drugs Aging 2008;25(12):1061-1075 [FREE Full text] [Medline: 19021304]
43. Makoul G, Arntson P, Schofield T. Health promotion in primary care: physician-patient communication and decision making about prescription medications. Soc Sci Med 1995 Nov;41(9):1241-1254. [Medline: 8545677]

Abbreviations

ANOVA: analysis of variance
DTC: direct-to-consumer
FDA: Food and Drug Administration

Edited by G Eysenbach; submitted 02.05.17; peer-reviewed by R Kornfield, L Vera; comments to author 15.08.17; revised version received 10.10.17; accepted 16.11.17; published 23.01.18.

Please cite as:
Sullivan HW, O'Donoghue AC, Gard Read J, Amoozegar JB, Aikin KJ, Rupert DJ
Testimonials and Informational Videos on Branded Prescription Drug Websites: Experimental Study to Assess Influence on Consumer Knowledge and Perceptions
J Med Internet Res 2018;20(1):e13
URL: http://www.jmir.org/2018/1/e13/
doi:10.2196/jmir.7959
PMID:29362205

Original Paper

“Why Do They Need to Check Me?” Patient Participation Through eHealth and the Doctor-Patient Relationship: Qualitative Study

Christiane Grünloh[1,2*], MSc, Tech Lic; Gunilla Myreteg[3*], PhD; Åsa Cajander[4*], PhD; Hanife Rexhepi[5*], MSc, Tech Lic

[1]School of Computer Science and Communication, KTH Royal Institute of Technology, Stockholm, Sweden
[2]Institute of Informatics, Technische Hochschule Köln, University of Applied Sciences, Gummersbach, Germany
[3]Department of Business Studies, Uppsala University, Uppsala, Sweden
[4]Department of Information Technology, Uppsala University, Uppsala, Sweden
[5]School of Informatics, University of Skövde, Skövde, Sweden
[*]all authors contributed equally

Corresponding Author:
Christiane Grünloh, MSc, Tech Lic
Institute of Informatics
Technische Hochschule Köln
University of Applied Sciences
Steinmüllerallee 1
Gummersbach, 51643
Germany
Phone: 49 2261 8196 6238
Email: christiane.gruenloh@th-koeln.de

Abstract

Background: Roles in the doctor-patient relationship are changing and patient participation in health care is increasingly emphasized. Electronic health (eHealth) services such as patient accessible electronic health records (PAEHRs) have been implemented to support patient participation. Little is known about practical use of PAEHR and its effect on roles of doctors and patients.

Objective: This qualitative study aimed to investigate how physicians view the idea of patient participation, in particular in relation to the PAEHR system. Hereby, the paper aims to contribute to a deeper understanding of physicians’ constructions of PAEHR, roles in the doctor-patient relationship, and levels and limits of involvement.

Methods: A total of 12 semistructured interviews were conducted with physicians in different fields. Interviews were transcribed, translated, and a theoretically informed thematic analysis was performed.

Results: Two important aspects were identified that are related to the doctor-patient relationship: *roles* and *involvement*. The physicians viewed their role as being the ones to take on the responsibility, determining treatment options, and to be someone who should be trusted. In relation to the patient’s role, lack of skills (technical or regarding medical jargon), motives to read, and patients’ characteristics were aspects identified in the interviews. Patients were often referred to as static entities disregarding their potential to develop skills and knowledge over time. *Involvement* captures aspects that support or hinder patients to take an active role in their care.

Conclusions: Literature of at least two decades suggests an overall agreement that the paternalistic approach in health care is inappropriate, and a collaborative process with patients should be adopted. Although the physicians in this study stated that they, in principle, were in favor of patient participation, the analysis found little support in their descriptions of their daily practice that participation is actualized. As seen from the results, paternalistic practices are still present, even if professionals might not be aware of this. This can create a conflict between patients who strive to become more informed and their questions being interpreted as signs of critique and mistrust toward the physician. We thus believe that the full potential of PAEHRs is not reached yet and argue that the concept of patient empowerment is problematic as it triggers an interpretation of “power” in health care as a zero-sum, which is not helpful for the maintenance of the *relationship* between the actors. Patient involvement is often discussed merely in relation to decision making; however, this study emphasizes the need to include also sensemaking and learning activities.

This would provide an alternative understanding of patients asking questions, not in terms of "monitoring the doctor" but to make sense of the situation.

(J Med Internet Res 2018;20(1):e11) doi:10.2196/jmir.8444

KEYWORDS

patient accessible electronic health records; medical records; personal health records; eHealth services for patients; patient portal; physicians; patient empowerment; patient participation; doctor-patient relationship

Introduction

Patient Participation and Electronic Health Technologies

Patient participation is advocated as a means to improve patient safety and is seen as a key component in the redesign of health care. There are, however, many barriers to patient participation that according to Longtin et al can be divided into (1) patient-related and (2) health care professional factors [1]. Among health care professionals, the main obstacles to patient participation are the desire to maintain control, lack of time, personal beliefs, and insufficient training in the patient-caregiver relationship [1].

It is argued that electronic health (eHealth) technologies have a tremendous potential to promote patient participation and improve health outcomes [2]. eHealth interventions recently started to focus on patients' rights to access their electronic health records (EHRs) over the Internet (eg, through patient portals). However, literature is not conclusive regarding the effects of making health records available for patients. Some studies have reported that patient accessible health records can generate anxiety or concerns [3], whereas others have concluded that having full access may decrease anxiety [4]. At the same time, it has been reported that health care professionals have been concerned about giving patients Web-based access to their health record (see eg, [5]). One of the concerns from physicians, as identified in a previous paper, was that patients would read their EHR with the purpose to control and monitor physicians [6]. Thus, they feared that patients would check on and monitor the physician's activities rather than adhering to the more "traditional" relationship: that physicians check on the patient, and not the other way around. These results indicate that we need to further explore the doctor-patient relationship in relation to eHealth interventions aiming at increasing patient involvement.

The purpose of this paper was to analyze and report in depth about how the interviewed physicians view the idea of *patient participation* in general and in relation to patient accessible electronic health records (PAEHRs). This is important to understand how they make sense of and assess the introduction of PAEHRs and eventually to explore the possible relationship between their concerns and the patients' abilities to become active partners in their care. The contribution of this paper is a deeper understanding regarding factors related to physicians' framing of PAEHR in relation to patient participation. Furthermore, the paper contributes with a critical discussion of the concept of patient empowerment as being problematic as it triggers an interpretation of power as zero-sum.

In our previous paper [6] from this study, we gave an overview of our whole dataset, which was thematically analyzed in relation to the physician's work environment. In this paper, we want to explore certain aspects of the data in more depth, which is in accordance to the method as presented by Braun and Clarke [7]. The detailed analysis conducted for this paper focused on patient participation and empowerment as an element of the patient-doctor relationship. The main research question driving this in-depth analysis was "How do physicians view the idea of patient participation in general and in particular in relation to patient accessible electronic health records (PAEHRs)?"

The Patient Portal and PAEHR

Patient portals are provider-tethered applications that allow patients to access, but not to control, certain health care information (eg, their EHR) and provide communication and administrative functions (eg, secure messaging, appointment booking, and prescription refill requests) [8]. In 2012, Region Uppsala in Sweden launched a Web-based patient portal to its 350,000 citizens as part of a large European Union project. The portal offers about 10 different eHealth Web services and aims to contribute to patient participation. Efforts to enhance participation through eHealth solutions have been emphasized in the National eHealth strategy of Sweden [9]. The provided eHealth services include, for example, PAEHRs, including the latest test results, appointment booking, following a referral, and a list of names of all health care professionals who have entered the EHR (so called "log list"). The PAEHR captures information from different EHR systems in all of Sweden, and the patient can read information from the primary care as well as hospital care. What exactly is shown in the system depends on (1) the EHR system the provider uses and (2) the region where the provider is located and thus whether and how the PAEHR system has been implemented. It is not possible for patients to edit the records; however, in Region Uppsala, they can comment on each of the professionals' notes. Patients access the portal using an e-ID or other secure log-in options. Initially, and at the time of the presented interview study, health care professionals had to sign or approve the medical notes for patients to access them within the first 2 weeks. This was later changed, in that patients in Uppsala now choose whether they only want to read signed notes or unsigned notes as well. Today, PAEHRs are provided in 19 out of 21 counties in Sweden and have more than 1,000,000 registered users [10].

Doctor-Patient Relationship

Various models of the relationship between the physician and the patient have been discussed in literature (see eg, [11-13]). The different models have been developed over time in accordance with new approaches to conducting health care, for example, shared decision making [12,14] and patient-centered approaches [15].

In the following, we will briefly describe the basic models of the doctor-patient relationship by Szasz and Hollender [16] and the models discussed by Emanuel and Emanuel [12]. Even though the approaches to conceptualize the relationship overlap on certain levels (eg, they both include a paternalistic model), we consider them both relevant for our discussion as they emphasize different aspects. In Szasz and Hollender, primarily the role of the patient is changing (from infant to adolescent child to adult) [16], whereas Emanuel and Emanuel outline a spectrum of possible roles for the physician [12].

For each of their three basic models of the doctor-patient relationship, Szasz and Hollender describe the roles of physician and patient and relate the relationship to a prototype [16]:

Activity-Passivity: This is the oldest model, according to Szasz and Hollender, in which the physician “does something to the patient” [16] who is a passive recipient. This model is suggested applicable when the patient is unable to actively contribute (eg, acute trauma). This relationship is compared with that of a parent and an infant [16].

Guidance-Cooperation: Szasz and Hollender outline that this model is employed in situations in which the patient is conscious, and both the patient and the physician are active [16]. However, the patient’s activity is rather to cooperate and “obey,” as the authors put it [16], without questioning, arguing, or disagreeing with the physician’s orders. This is explained in that the patient places the physician in a position of power because the latter possess medical knowledge that the patient is lacking. The prototype of this model is the relationship between a parent and the (adolescent) child [16].

Mutual participation: In this model, the physician has the role to help the patient help himself, who as an equal partner in this relationship uses expert help [16]. Accordingly, the prototype of this model is the relationship between two adults. The authors explicitly refer to the management of chronic illnesses as clinical application, where “patient’s own experiences provide reliable and important clues for therapy” and where the treatment is often carried out by the patient [16].

The analogy of the relationship between physician-patient and parent-child has been made also by others. The very term of the *paternalistic model* already entails the reference to a “father.” Katz describes that patients may display childlikeness, which is “triggered not only by pain, fears, illness, and memories but also by how physicians view and respond to patients” [17]. Furthermore, by viewing them “too much as needy children, physicians disregard the fact that patients are adults as well” [17] who have certain needs such as wanting to be informed and involved.

In the four models of the “physician-patient relationship” outlined by Emanuel and Emanuel, the physician’s role varies between a *guardian* (paternalistic model), a *counselor or adviser* (interpretive model), a *friend or teacher* (deliberative model), and a *technical expert* (informative model) [12]. The paternalistic model, in which the physician determines what is best for the patient, leaves little room for the patient to participate [13] and should today merely be justified during emergencies [12]. In situations other than emergencies, patients’ participation is essential because although physicians might possess more medical knowledge, patients know more about their own needs [17]. In addition, some patients possess in-depth knowledge of their condition, which may even exceed that of the (more or less prepared) health care professional [18].

What emerges from the above is a quite disorderly view of what the doctor may or should represent to a patient and what kind of relationship between patient-doctor would need to be established. Solitary decision making by physicians has a long tradition in medicine and obscures the uncertainty of medical knowledge, which, as assumed by physicians, would lead to anxiety and confusion if brought to the patient’s attention [17].

Patient Participation

As the changing roles in the doctor-patient relationship suggest, patient participation in health care, including decision making, is increasingly emphasized. Patient empowerment has been described as the attempt to increase the patient’s capacity to think critically and make autonomous, informed decisions [19]. Patient empowerment is surrounded by many other concepts such as engagement, enablement, participation, involvement, and activation [20]. In an attempt to clarify boundaries and relationships between these concepts, Fumagalli et al developed a concept map in which they combined the key definitions into “Patient empowerment is the acquisition of motivation (self-awareness and attitude through engagement) and ability (skills and knowledge through enablement) that patients might use to be involved or participate in decision-making, thus creating an opportunity for higher levels of power in their relationship with professionals” [20].

The question remains what involvement or participation in *decision making* means in practice. In what kind of decisions are patients involved and able to participate? Deber distinguishes between two dimensions of choice: *problem solving* (the search for the solution to a problem) and *decision making* (the choice being made from several alternatives) [21]. It has been noted that *problem-solving situations* require some level of medical knowledge and thus, do not present themselves well to patient participation, whereas certain *decision-making situations* require the patient to analyze and determine the value of potential outcomes [1]. This distinction is very relevant when it comes to the question whether patients “want to be involved” or would rather “leave it to the doctor to decide.” Research showed that patients are quite capable to discern between these situations [1,22]. In a study by Thompson, patients’ desire to be involved was much higher regarding decisions that do *not* require medical knowledge but that have lifestyle implications and where attitudes and values are likely to be important factors [22].

Similar to the doctor-patient relationship, different models of involvement and participation in consultation exist, which reflect various levels of patient power [23]. With an increasing level of patient power, the levels of professional-determined involvement are (0) exclusion, (1) information-giving, (2) consultation, (3) professional-as-agent, and (4) informed decision-making [23].

Despite a vast amount of research in this area, an in-depth understanding of the impact of PAEHR in relation to patient participation and the doctor-patient relationship is still lacking. To reach the aim to understand the physicians' views of patient participation and PAEHRs' possible effects on this, this paper adopted the models of the doctor-patient relationship and the various levels of involvement from previous research.

Methods

Interview Content and Data Collection

Semistructured interviews were conducted in the summer of 2013, about 6 months after the PAEHR service was launched. Twelve physicians were interviewed by three different researchers. All researchers used the same template for questions 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 one, which was carried out by email. On average the interviews lasted 1 hour.

Participants

As reported in the previous paper [6], getting 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 and making use of mailing lists. The project nevertheless succeeded in getting a positive response from physicians in four different specialties: orthopedics, oncology, emergency medicine, and internal medicine. The characteristics of the interviewed physicians (N=12) can be found in Table 1.

Analysis

All interviews were transcribed, translated, and repeatedly read by all authors. This paper is the second reporting of the study. For the first paper [6], a thematic analysis [24] was conducted in which the whole dataset was coded, also known as *complete coding* [7]. For this paper, a selection of data was used for an in-depth thematic analysis, which consisted of data that previously was coded "patient empowerment." The selection comprised nearly 50 pages of interview excerpts, which again were thoroughly and repeatedly read through, jointly discussed, coded, 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 [24]. Part of the analysis process was also the iterative development of a thematic map. The iterative process with several rereadings, discussions, and thematic descriptions was carried out with the aim to achieve trustworthiness in the research process. At the same time, as the potential themes were identified and reviewed, the authors read and discussed the wider literature to build on established concepts and in particular their distinctions (eg, patient-desired vs professional-determined involvement [23], overlapping meanings of concepts such as empowerment, engagement, enablement, participation, involvement, and activation [20]). The quotes used in this paper have been slightly edited to be more readable.

Table 1. An overview of the interviewees (N=12).

Characteristics	Number of interviewees
Specialty, n	
Orthopedics (Ortho)	5
Oncology (Onco)	3
Emergency medicine (EM)	2
Internal medicine (IM)	2
Gender, n	
Female	5
Male	7
Work experience (years), mean (range)	14 (2-30)

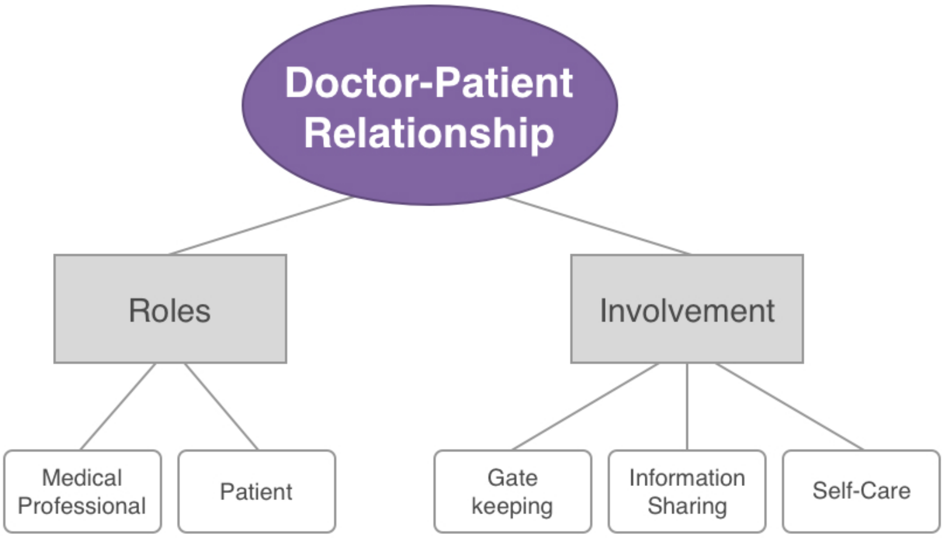
Results

Overarching Themes

The thematic analysis of the dataset resulted in the identification of the overarching theme *doctor-patient relationship* (Figure 1). The *doctor-patient relationship* captures two important aspects that were discussed during the interviews with physicians: the roles that are involved in this relationship (ie, the medical professional and the patient) and in what way patients can contribute in this relationship in terms of involvement (related to concepts such as gatekeeping, information sharing, and self-care).

In the following sections, the themes and subthemes will be described and discussed in relation to the wider literature.

Figure 1. Identified themes from interviews with physicians related to patient participation.



Roles: Medical Professional

The theme *Roles: medical professional* captures the way the physicians talk about their own role in relation to patient participation. When asked directly about patient participation and what this means to them and to their role, many physicians expressed that they were not generally opposed to this. They rather expressed that patient participation is important and that they are in favor of it. However, one common understanding of participation found in the interviews was that patients' experiences need to be understood for the physician to assess them, and then the patient needs to be convinced about what treatment is the best, as exemplified in the following quote:

> *In order to get good treatment results it is really important for us to understand what problems the patient is experiencing so I know how I can respond to it and assess it. And then when we think we have understood what is wrong, then we need to discuss it with the patient, we must get the patient on board and make him or her also believe that this seems reasonable, that this is the problem.* [EM-1]

Moreover, on several occasions, the physicians indicated that for them patient participation is about presenting options from which the patients may choose. For example, when being asked what it would mean if the participation would increase, this physician responded:

> *I do not really think it [patients reading PAEHR] would mean so much difference at all, because already today you ask the patients how they want it, if you have two equivalent alternatives to present it is what the patient thinks that determines it.* [Ortho-4]

In our interpretation, the above presented ways of involving patients depicts the role of the physician akin to the *professional-as-agent*. According to Thompson, the professional-as-agent possesses "the technical expertise, but patient preferences are incorporated into their decision-making" [23]. As Thompson points out, the incorporation necessitates some prior dialogue and considers this as level 3 of patient involvement, which is the second highest level [23]. However, it can be discussed what this means in terms of patient participation, as in the quote from EM-1, the term getting the patient "on board" might be understood as getting the patient on board of the doctor's boat and not the other way around.

Deber even argues that the values assigned to potential outcomes are not relevant to problem solving (ie, searching for the solution to a problem), and thus, patient participation is not necessary to identify alternatives and to estimate their possible outcome [21]. However, if the preferences and values do *not* play a role and the physicians determine the "equivalent" alternatives *without* a dialogue, we claim that this cannot be considered as patients participating in decision making.

Some physicians claim that the physicians need to have the final say when deciding upon a treatment and that there is a limit to patient participation in relation to science, as exemplified in this quote:

> *There is some sort of limit to how much should the patient [be involved]. Where is the line between being involved and to decide? We do have some treatments that we claim to be better than others, but some patients come with a belief that, well, they have read or heard that a friend got this other treatment and think it is better, and /.../ I think that the doctor should have the final say in the case if you are basing your opinion on science.* [Ortho-3]

This comment shows the opinion that evidence-based decisions should be prioritized. According to Williamson, ethical challenges are raised when no *clear* evidence-based solutions are available [25]. In these cases, the decisions have to be supported by patient's values and preferences. However, few physicians in our interviews acknowledged the ethical challenges. More often it was a matter of "convincing" the patient of the best treatment options regardless of individual preferences. We interpret this way of making decisions as an example representing the physician in a paternalistic role.

Another aspect of the medical professional's role was the question of responsibility and trust. This was discussed in relation to patients accessing their EHR. Taking care of patients was seen as the physician's responsibility. As long as people trust their physician, there should be no need for patients to read their records, according to some of the physicians. Most physicians also believed that the patients already know what they need to know without reading their EHR, as stated in the following quotes:

> *We are trying to do the best for all patients. We are the ones who take responsibility for complications and everything so they're going to try to enjoy life and not sit in front of a computer and check test results and devote time to it.* [Onco-1]
>
> *I believe that most patients feel that they know what they need to know, and have the influence according to their level of knowledge, or how to put it. It's a bit like if I leave the car to the mechanic, I do not expect that I'll know exactly what they will do, but I'm happy if they fix it—sort of.* [Ortho-2]

Furthermore, patients who intend to be more involved in terms of asking questions were interpreted by one physician as a sign of mistrust. This mistrust was also seen as a recent phenomenon related to physicians' status in society and a lack of respect for the physician's high education. One physician stated as follows:

> *And there is a mistrust in this that bothers me very, very much. A mistrust for what I suggest. /.../ What has happened to the old image of the doctor who was very good and "now you are going to meet the doctor" and it was a person with a high status. I don't have to be seen as some demigod, but I want to be respected for the education that I have. It takes years to become a doctor, and even more years to become an orthopaedic, and all these years of education now count as nothing, because the patient should choose now. And that mistrust bothers me. Do I feel that mistrust, then I usually say to the patient that this will not be good. I will transfer you to another doctor.* [Ortho-1]

We interpret that the idea that patients should give physicians full responsibility and trust them to know what is best for the patient is closer related to paternalism than to a partnership in care. In paternalism, "the professional knows best and patient involvement is limited to being given information or giving consent" [23]. In summary, the way the physicians talk about their role emphasizes that they are supposed to take on the responsibility, determine the alternative options from which—if equivalent—the patient may choose from, and the patient as such should not be involved in this process but only trust the physician's judgment. In this sense, there is no risk of "abandoning the patient," which was emphasized as being a risk when the autonomy is too high (compare to [25]). This high level of autonomy would correspond to involvement level 4—informed decision making—where the patient makes the final decision after the technical expertise is transferred to him or her [23]. However, we would argue that in the physician's construction outlined previously, the patient is still left with a rather passive role, which is not in line with attempts to establish a partnership among equals, nor in line with the idea that patients should read their EHR.

Roles: Patient

The theme *Roles: patient* captures the way physicians talked about the patients in terms of their skills (or lack thereof), motives to read the EHR, and **c**haracteristics.

Skills

The physicians discussed certain skills that are needed, but which patients might lack, to *benefit* from accessing their EHR. For elderly people, some physicians assumed that they do not have that skill as they are not that familiar with the technology needed. Thus, according to physicians, elderly people are probably not that interested in reading their EHR, which is illustrated in this quote:

> *I most care for the elderly and then, they usually do not have an interest. They need to have e-identification, for example, and Internet and stuff, and they usually don't have that. So, actually it is more relatives like me, who want medical record copies, but not the patient himself/herself.* [IM-2]

Although the assumed lack of technological skills was related to elderly adults only, the lack of medical knowledge was discussed in relation to all patients. Because the content of the EHR must contain medical terms, it was assumed that patients would not understand it or even misunderstand its content. One physician stated the following:

> *The record is not a means of communication with the patient, the record is a tool and therefore it must contain medical terms that the layman does not understand for it to be an effective communication tool among doctors and other health professionals.* [Ortho-2]

In relation to the content, it was not only the jargon that was assumed to be difficult to understand by the patients but also to determine which information is *relevant* or *important*. Almost all physicians discussed the complexity of the records and that they are difficult to interpret and evaluate (even for them). The records must therefore be filtered and interpreted for the patients to help them understand. One example often mentioned by the physicians in the interviews was what the particular lab results mean for patients. Without the physician's guidance and support, patients would focus on details and not understand the real meaning of lab results. Moreover, physicians claimed that patients refer to details in the record that are of no importance to the physicians, as exemplified by the following quotes:

> *There are also a lot [of cases] where a lab value or something is outside the reference range but when it does not mean anything, which I think the average citizen does not understand: "But it is outside the reference range, so surely something must be wrong."* [Ortho-2]
>
> *And then I have had patients who several times have printed their record and then they say: "You write here that I had pains in two weeks, but the truth is that I did have pain for three weeks." And then they talk about bitty details in the record and talk about all the details.* [Ortho-1]

These comments exemplify that the professionals regarded certain aspects or wordings as not as important or relevant as their patients did. Whether the particularities in the records (2 vs 3 weeks as in the quote above from Ortho-1) are relevant might be questioned. It is understandable though that patients are interested in having errors corrected, which has also been reported by Esch et al [26]. However, Rexhepi et al reported that few patients actually ask for errors to be corrected [4]. In addition, we are faced with the question of whether the patient is static or not. Related to the view of patients not being able to understand lab values, patients may learn over time to interpret the results, especially if the same tests are taken on a regular basis.

One way for physicians to handle that patients have difficulties interpreting the information has previously, before PAEHR, been to channel and select the pieces of information they give to the patient. This strategy, however, is now challenged when patients can access their record and tests results directly. This change is illustrated in the following quote:

> *If, for example, you are examining a patient, you don't give them all answers one by one as they come in. Instead you make a plan and decide how the patient is and what sort of information they can take. Then it is difficult that they can just go and read it all by themselves without having any idea what that means mostly, really, so it's not good.* [IM-2]

We interpret that this way of presenting patients with selected information is a type of interaction that follows on the paternalistic model of doctor-patient interaction [12]. This model describes how the physician provides the patient with "selected information" to "encourage the patient to consent" [12].

The physicians were aware of, however, that today patients are also looking on the Web for further information. In the interviews, this was regarded as both something positive and negative. Although searching on the Web might clarify questions for patients and increase their knowledge, some physicians were concerned that patients might not be equipped to interpret all this information, as exemplified by the following quotes:

> *They sort of clarify, sort out the question marks. [Using Internet resources] may provide tools to find out more. If I say that they have the disease X, they can go home and google or go to the library or whatever they do to acquire knowledge and learn more about it.* [Ortho-3]
>
> *You can search [online] and many do, both patients and parents. So, you go out and search Google for various treatments. And seek their own information, and there is the problem to be able to evaluate the information you find because a lot is not scientific or quality assured.* [Ortho-4]
>
> *There is nothing that is worse than a patient that has read things on the Internet and says "I absolutely do not have heel spur." Then I say "you do have heel spur." Then they go: "I read on the Internet and I do not have heel spur. I read on the Internet and don't have everything that it says there." But then I say "that the other 19 things on the list out of 20 was correct, so you have heel spur. From experience, I know that if you have 19 out of 20 then you do have heel spur." "Well, that one was not correct." I have these problems all the time, and it is just because the Internet exists. Very tough and energy consuming! This is because Internet has no control. Internet contains anything at all. If you enter the wrong page, and read the wrong thing where someone who is not serious writes, then you can get the wrong information.* [Ortho-1]

Motives

During the interviews, the physicians mentioned some possible motives to (or the lack thereof) why patients might want to read their EHR. Lacking interest in reading was connected to elderly adults who are satisfied with what they know. The elderly do not "want to know everything," according to physicians. If a physician perceives that patients today put less trust in the role of the doctor (see above), this was assumed to be a motive for the patient to read the EHR. Thus, some possible motives to read were negatively associated by the physicians, for example, regarding patients to act as the police or someone who check on the doctor. One physician stated the following:

> *I am very afraid of misunderstandings. Misunderstandings, and mistrust, and some already say that. Do you [the patient] think that I do what is worse for you, or what do people think? Why do they need to check me? If you have that perspective and there is a misunderstanding, then everything can happen. SO WRONG!* [Ortho-1]

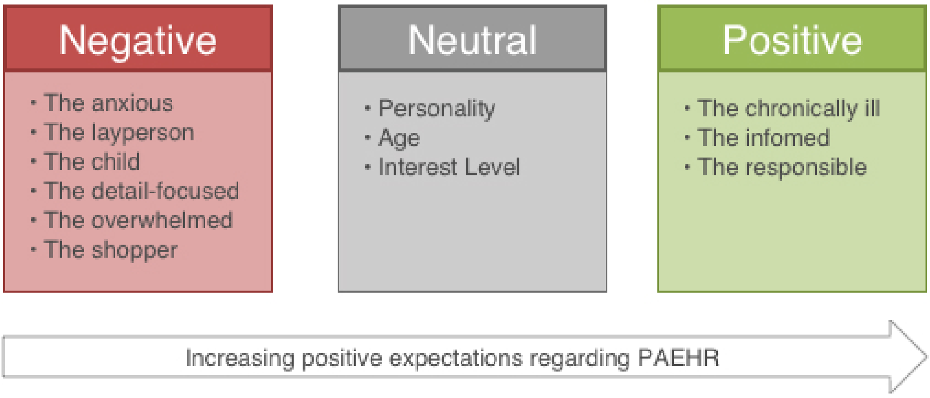
A few physicians mentioned that some patients need their health records for an insurance claim or because they want to share the record with family members who work in health care. Several motives in relation to a medical interest were also mentioned: patients can have an eye on the progress on the Web as well, use the EHR as a memory aid about what has been discussed during the visit, or read the test results before their next planned visit.

Characteristics

The physicians' expectations ranged from rather negative to neutral to rather positive in relation to patients accessing their EHR (Figure 2). A negative or a positive expectation (red or green box in Figure 2) could be related to a certain stereotype of a patient, which physicians used as examples. Where the expectations were rather neutral, the physicians would not describe a specific "type," but instead they discussed the possible characteristics on which the outcome might *depend* (grey box in Figure 2) such as personality, age, or whether they are interested in the first place.

If physicians were expecting rather *negative* behavior from patients accessing their EHR, patients were described as anxious, detail-focused, overwhelmed (ie, who need guidance), would possibly "shop around," or as laypersons. These types of patients would demand that the physician either would need to calm the patient, to guide him or her, or to explain again the situation.

Figure 2. Characteristics of patients that were associated with negative, neutral, positive expectations toward patient accessible electronic health records (PAEHRs).



This extra work would arise because of the existence of PAEHRs, according to these physicians. In addition, certain activities (eg, guiding the patient through their results) would be impeded when patients read their records on their own.

In some instances, physicians commented on the patient's access in a rather *neutral* way, while acknowledging that certain aspects depend on the patient. For example, whether patients are involved directly would depend on their personality, their general interest, or their age (ie, the elderly are perceived as less likely to be actively involved or to have Internet access). Thus, in these instances the physicians did not have fixed expectations and did not describe a specific type of patient.

Few physicians commented in a *positive* way about certain patients accessing their EHR and thus, possibly benefiting from this eHealth service. Here, the physicians described certain types of patients, namely those who are chronically ill and thus "extremely interested in their healthcare" [Ortho-5] or patients who are well-informed and read up. A last group of patients was described as taking responsibility for possible consequences of their actions (eg, if they read their record outside of office hours and thus face the consequence that they cannot immediately contact health care to ask questions).

It is interesting, that none of these physicians discussed how a patient could possibly develop from the left category (negative) into the middle (neutral) or right category (positive). People are not a static entity but are able to develop skills and knowledge over time. It could well be that the chronically ill patients and those who are well-informed and read up were not always like this but developed over time. However, even well-informed patients might face new situations where they lack knowledge, have to adapt, and start a new learning process.

One explanation could be that the interviewed physicians from the hospital might not treat the individual patient on several occasions. In other words, lacking a continuity of care makes it difficult for physicians to develop a long-term relationship with a patient and thus, recognize the potential developments.

Involvement: Gatekeeping

Involvement captures aspects that support or hinder patients to contribute and take an active role in the doctor-patient relationship. Involvement includes themes such as gatekeeping, information sharing, and self-care.

The interview questions centered around the patient portal in general and in particular in relation to the eHealth service, which gives patients access to their EHRs. It is thus not surprising that many aspects that touched upon involvement of patients in their care were related to information sharing between doctor and patient. One of the introductory questions, however, explicitly addressed physicians' thoughts on patient involvement and what it could mean for them if the participation would increase.

Some physicians were negative to patients reading their EHR by themselves and would rather prefer physicians to be the contact person for patients, as exemplified by the following quotes:

> *Doctors are honest and are telling the truth, why should they [patients] get access to the record, they can surely come to visit and discuss possibilities and explanations.* [Onco-1]
>
> *It is not that we want to hide something or, but there is a world of its own and we have gone through a long training to handle it in the best way and then you have to also let us do the work without checking everything and interfering.* [IM-2]
>
> *I think it's inhumane to patients, who of course then go in [the EHR] and look, because they want to know, and they think "I can handle it."* [Onco-3]

As aforementioned, at the time of the interviews, patients could only see those records that were signed by the physicians or were older than 2 weeks. The necessity to sign notes could at this time be used by physicians to prevent patients from accessing information as is described in the following quote from an oncologist:

> *I will no longer sign test results in the same way as I did. If there is any progress then I will not sign since the patient can go in and read the answer before I have had time to call and tell them.* [Onco-1]

The same physician also considered writing more vague descriptions and wait with the specific details until they were discussed with the patient. Both, changing the way of writing and thus limiting the information provided and not signing the notes to prevent immediate access, can be interpreted as "covert ways to remain in power," as described by Longtin et al [1].

Involvement: Information Sharing

In this section, bilateral information sharing between the health care system or physician and the patient is discussed. On the one hand, the health care system is sharing information with the patient by giving them access to their EHR through the patient portal. On the other hand, patients are also sharing information by *not* blocking parts of their record. Patients are entitled by law to block certain parts or all of their medical records and then the records would not be accessible for health care professionals unless there is an emergency situation. The interviews addressed also the prospect of patients blocking information; thus, the discussion also considered information that is shared (or not shared) with physicians.

Physicians were not critical about giving information to patients in general, on the contrary, they considered this as important, as exemplified by the following quotes:

> *The patient must be well informed about their disease and what we might plan to do, what are the opportunities, what the prognosis is, treatments, that they will have an influence over what they will be going through, and for them to be involved so they must have been properly informed.* [Ortho-2]

> *It is very, very important that we give information why we do different things, what is happening around you. "Why should I not eat, why should I shower with a special soap?" /…/ Before the doctor told the patient what to do and the patient did it without really understanding, and nowadays that does not happen. Today people require to understand why and I think that is reasonable.* [Ortho-1]

In the quotes above, the physicians talk about the necessity for patients to understand what is happening around them and exemplify this with the reasons behind certain *instructions* that the patient is supposed to follow. However, patients may ask questions about aspects that the physician may not deem to be relevant at that time. One physician stated the following:

> *Or they bring a record that some other doctor has written, and then they say that "I don't understand these words. Could you explain them to me?" Then I have to take time from my schedule to explain this to them. When it wasn't even me that had written them to start with.* [Ortho-1]

Although these kinds of questions may be perceived as a burden, we believe that they might nevertheless help the patient to make sense of the situation. This sensemaking may in turn be necessary to be able to ask further questions to reach an understanding.

On several occasions, the physicians mentioned that the patient has the right to see the information; even though some were critical about *when* and *how* this information was shared. One physician stated the following:

> *The patient has the right to know /.../, they have the right to read the record and get to know how we have written about it, and it's not something secret, I mean we are talking to the patient about what is happening and we explain everything.* [Onco-1]

Some physicians reflected on how information is shared with the patient in general (ie, without information and communications technology [ICT]) and that this process is also not perfect, as exemplified by the following quotes:

> *I sometimes think that there is a certain lack of written information about various diseases because, uh, at a clinic visit, the patient is often stressed and /.../ it is very much information to take in, uh, so I think that many would have needed to read something a bit more structured after the visits, for example.* [Ortho-2]

> *It is difficult as well to memorize everything that is done and all samples in a clinic visit, so now you can sit at home in peace and quiet and read the record.* [Ortho-5]

> *If you put demands on the patient to rest their foot and stay in bed for two weeks, then they need to understand why they need to do that. /…/ We are not very good at giving that information, and that information should be on the Internet. There is where people look today.* [Ortho-1]

The responses indicate opportunities for ICT to improve information sharing with the patients, for example, by giving them more structured information. However, it was strongly questioned whether giving patients access to their EHR would solve this lack of information, as it was assumed that this would worry patients, and they would not understand more from what they have read. In the quote below, Ortho-1 expresses that there are better ways to inform patients than to give them access to their EHR:

> *I think that you solve the wrong problem. Patients are not informed; they feel non-informed. And I agree with that. Worried, they don't understand, they are not doctors themselves. This is solved by them reading their medical record? It is the wrong solution to the problem! Then they become more worried, and more upset, and don't understand. All this worry should be met in a better way with better information and better general information. Better information about what will happen during your surgery, what does this word mean, what will happen now.* [Ortho-1]

Physicians sharing specific information with patients can be interpreted as level 1 of patient involvement, which includes "simply giving them information considered necessary by professionals" [23]. Considering that level 0 is actually the level of *excluding* patients [23], we regard the level of involvement as not very advanced on level 1. Thus, although physicians were positive toward sharing information on a more general level, there is room for improvement on the way to a partnership. In particular, information sharing should possibly go beyond what professionals consider necessary and also reflect the patient's need for information.

Most physicians were, however, critical about patients' possibility to block certain information in their EHR and by this, denying the professionals access. Although few agreed that there might be instances where the patient might want to block sensitive data, they considered this as a potential threat to giving patients proper care. However, most physicians acknowledged that patients might want to exercise their right to block information but would then also have to take responsibility for it, as exemplified in the quote below:

> *Basically, it's right that you determine the course of your own body and what others should know about it. But on the other hand, I fear that it might get out of control and that it can eventually become a danger to a patient and information which for me can be very important that I cannot access can then lead to me giving wrong or even dangerous care to someone. I can understand it, but for my treatment of the patient it is a risk that it is something bad. So overall, I think it's unfortunate, but at the same time, it is perhaps something you have to accept that people decide over themselves.* [Ortho-3]

It was seen positively, however, if the system would indicate whether anything was blocked, so that they could engage in a conversation with a patient on whether they want to share this information with them, as expressed by EM-1 in the quote below:

> *That would be a bit better. Then I can see that something is hidden, and I can ask the patient: "Something has been blocked here, would you like to tell me what is says there?" And you might be able to solve it that way. If this is the way they have chosen to go, this might be a better solution.* [EM-1]

The physician's concerns of patients blocking them from information is interesting, in that it can be related to the idea that only the professionals are in the position to determine whether a piece of information is relevant or not. This refers not only to what should be shared with the patient but also what information should be available to the professional and what kind of questions is relevant to discuss. It seems that physicians have the motto "the more information the better" when it comes to themselves and their work. However, the same motto does not seem to apply to patients, as it is assumed they would be overwhelmed and focus on irrelevant details.

Involvement: Self-Care

The theme *self-care* captures discussions around patients who have the ability to care for themselves and perform activities necessary to achieve, maintain, or promote optimal health, as has been defined by Richard and Shea ([27], as cited in [28]). Although the physicians tended to discuss their own responsibility in taking care of the patients, on few occasions it was mentioned that patients can also engage in self-care and might benefit from technological support. The quote from EM-1 is one such example:

> *If you think of the big picture, I think it is a lot down to the available information, that the patient is worried about some disease and they can get information from reliable sources, not just googling for it, get help to make their own judgment about what they believe so that they may be able to engage in self-care or they could actually wait a few days and make an appointment at the local surgery, they do not need to come to the emergency ward. Then they are more involved in their own healthcare* [EM-1]

As discussed previously, only few physicians mentioned patients with chronic conditions, although these patients often engage in self-care and perform many health-related tasks outside the doctor's office. Although some physicians talked about this patient group and the potential benefits for them to access their EHR, others considered this group as rather marginal, as exemplified in the quote below:

> *There could possibly be some chronically ill patient who is extremely interested in their healthcare and who review and evaluate the information in their medical records. They would possibly be able to control their disease better, but it is such a very small group, and the group gets this information in any case. One can only ask that you be kind and print test results on paper when you're at the doctor or whatever it may be, you can print the X-ray and so on. So, I do not think it will change, the possibility has been for the small, small group it concerns.* [Ortho-5]

Patients being encouraged to change their behavior can also be related to self-care, in that they have to take action themselves. The following comment exemplifies the idea that to emphasize the need for a behavior change, physicians might use the EHR to add comments for the patients to read, for instance "quit smoking," as exemplified in the quote below:

> *But it's a difference if it comes to, for example, smoking cessation, then you need a lot of participation that the patient himself realizes: "oh well, I also have to do something." You try to help the patient get medical care or a group or stuff like that, then we can try to help the patients to make the step themselves.* [IM-2]

A few physicians discussed the possibility that patients could use their EHR as a memory aid to review instructions, which may then lead them to better follow those, as exemplified in the following quote:

> *It would surely be that the patient gets easier opportunity to review what was said at such a clinic visit, to remember more all instructions maybe, or together with the information you had, because you know that it is difficult to absorb all the information during a visit so, so that it is clear that there can be an advantage to have it as repetition.* [Ortho-2]

It has been emphasized that patient empowerment is the "antithesis of compliance" [19], in that empowerment-based interventions help patients to "think critically and make informed decisions" [19]. Thus, given this view, reaching a better compliance should not be the main objective for giving patients access to their records. However, reviewing the information discussed during a visit can support patients' self-management, which in person-centered care is "another route toward greater participation" [25].

Discussion

Limitations

A limitation inherent in using interviews as a method for data collection is that what participants report may differ from what they actually do. In addition, some statements in the interviews related to patients reading their EHR were rather expectations than actual experiences. For the coders, it was partly unclear whether the participant reported on an actual experience. However, the analysis was focused on the descriptions of their daily practices and how they explained their concerns in depth and related the constructions to concepts found in the wider literature (eg, models of doctor-patient relationship).

As the interviews took place only a few months after the launch of the system, it is possible that attitudes might have changed. Although a survey conducted in another region in Sweden (Region Skåne) about 2 years after the introduction of the service suggests that this might not be the case [29]; follow-up research in Region Uppsala is needed and already in the works.

As the interviewed physicians had different specialties and worked at the hospital where a continuity of care might be lacking, results may be different with general practitioners (GPs) or physicians who are able to develop a long-term relationship with their patients. In addition, limits of participation as seen by physicians may differ whether they are specialists (eg, oncologists or orthopedists) or a GP. Further research is needed to investigate whether the different types of the relationship (eg, short term or long term) or the gravity of information or decisions they are dealing with influences the physicians' attitude toward patient participation.

Conclusions

The conclusions are related to two areas: (1) the doctor-patient relationship and the possibilities to use PAEHR as a tool for patient participation and (2) the use of “patient empowerment” as a problematic concept.

The Doctor-Patient Relationship and PAEHR as a Tool for Patient Participation

Already in the nineties, it was argued that unless there is an emergency situation, the paternalistic model of a doctor-patient relationship is not beneficial. The main reason is that this model assumes that patients and physicians share similar values and views, which is an assumption that may be incorrect [12]. Instead, shared decision making as a collaborative process in the medical encounter has been advocated [12,14]. Although the physicians in this study answered that they, in principle, were in favor of patient participation, the analysis found little support in their descriptions of their daily practice that participation is actualized. On the contrary, there were several signs of paternalism. This interpretation was further strengthened by the expectations that physicians often held regarding the characteristics of their patients (eg, as being unable to understand and as being worried and anxious).

The paternalistic model was also mirrored in the interviews, in that physicians described that patients only should be provided with pieces of information that they might be able to take in at a certain point in time. The physicians reacted strongly when patients could read the results before they had finished their process (ie, the investigations). In addition, physicians also criticized that they would have to spend extra time explaining and/or discussing information patients found themselves, which they would not have to do otherwise. It was also mentioned repeatedly that patients should trust the physicians, who are not working against the needs of the patients but have the best for them in mind. This exemplifies an assumption of shared objectives, which is an important dimension of the paternalistic model.

At the same time, physicians tended to view patients' questions as signs of critique and mistrust. In our interpretation, this is an interesting struggle between the need for a “patient that understands” and the negative reactions that questions may evoke. We defend the possibility for patients to pose questions not only regarding instructions but also regarding other “objective facts” because to the patient, this is a process of sensemaking that is not necessarily of the same kind as to the professional. Making it visible for physicians which of their patients read their EHR might alter the doctor-patient meeting. Knowledge of whether the patient has read their EHR might help physicians to open up the discussion and hereby, create an open atmosphere where patients feel comfortable to talk about what was said in the record. This could be an opportunity to increase patient involvement to higher levels than was expressed in the interviews, which we attributed to level 1 (information giving) [23]. However, making it visible to the physicians who interpret patients reading as a sign of mistrust might affect patients so that they read less.

Low health literacy and the lack of knowledge of the subject is one obstacle to patient participation [1]. Although it still is the case that some patients do not want to be involved more than necessary, we believe that by opening up and inviting patients to ask questions, they will be more willing to participate, which could improve their health literacy. Considering that people develop skills and knowledge over time, reading their own record through a patient portal might enable patients to ask more questions.

As seen from the results, paternalistic practices are still present even if professionals might not be aware of this. PAEHRs make it much easier for patients to have access to their data without having to ask for permission. Even if the professional motto “the more information the better” could not be realistically applied to a patient, by the support of PAEHR we can at least avoid the patient being caught in a restricted situation of “less is more.”

We see PAEHR as a tool that opens a path for patients in that it is more difficult for professionals to exclude them. This is, however, only the first step. Our analysis highlights a clash between the principles underlying physicians' practices and PAEHR, namely paternalism versus participation. This is an important finding. We believe the potential of PAEHRs is not reached yet, in that they could support not only the communication but might also support a change of the doctor-patient relationship toward one among equals.

"Patient Empowerment" Is a Problematic Concept in Academia and in Practice

Following from our analysis, we believe that a concept like "patient empowerment" is not helpful, in that it implies a loss of power and control on the part of the professionals. Health care is traditionally hierarchical and, especially in the paternalistic model, professionals have the authority. Emphasizing the *empowerment* of patients triggers an interpretation of "power" in health care as a zero-sum, meaning that a gain of power for one side (ie, the patient) entails a corresponding loss for the other side (ie, the professional). Although we are not opposed to giving patients more power, we consider this interpretation as not helpful for the maintenance of the *relationship* between both actors. In addition, we do not consider power in health care as a zero-sum situation in which the patient in the end will "take over" from the professional who ends up being a mere the technician who gives an advice (compared to the "engineering model", [13]). We perceive, however, that patients might have other reasons to follow up on the test results than the professionals might have. Although professionals read the information in the record to diagnose the patient and plan the treatment, the patient might feel the need to read the record to follow up on what is happening (eg, reading the log list to check whether anything is happening), to start a sensemaking process, to prepare for the next meeting, or maybe even to process what is happening to them. The patient empowerment definition by Fumagalli et al includes the acquisition of motivation and ability to be involved; however, it focuses merely on *decision making* [20]. Relating to the above discussion, we prefer a term like *participation* over *empowerment* and would extend this to include also sensemaking and learning activities. Sensemaking and learning are ongoing activities throughout the health care pathway. Even patients with chronic conditions encounter new situations through new symptoms or relapse, in which the sensemaking and learning process might start yet again.

A commonality between the presented models of the doctor-patient relationship in the background of this paper [12,16] is that patients are described as being static and unchanged. However, patients can develop over time from an inexperienced "childlike" patient into an "expert" patient or the other way around. Albeit not easy, professional attitudes and practices have to be reexamined, including the view of patients as static and unchanged. Other things that need to be reconsidered are the importance of hiding uncertainties from patients, the need to appear authoritative, and to view patients as not sufficiently competent to participate in decision making (compare to [17]). This can be related to health care professionals who have to "unlearn being in control" when patients are becoming more empowered [30].

In this collaborative setting, PAEHRs can contribute to the development in the doctor-patient relationship, in that it opens a way for patients to play an active role and makes it more difficult for physicians to maintain a strategy that potentially exclude patients. Thus, it is likely that the relationship between doctor and patient is changing; however, the question is *when* a transformation will take place and *how* the relationship will develop over time. One might hope that by being able to read one's records, the involvement increases and hereby, the distance between patient and physician decreases toward a relationship among equals. However, there is a risk that the future still remains "the silent world of doctor and patient" [17], where professionals do not ask and patients do not tell whether they read their records. Further research is needed on how patients make use of the record and whether professionals are today still unaware whether their patients read. In addition, the question is also how both actors may be able to support each other. For instance, professionals may be able to support patients to make sense of what is currently happening to them, and patients can help professionals understand what they are dealing with on a day-to-day basis. In that sense, we believe that eHealth does not need to be a "power struggle" in the doctor-patient relationship but can potentially help both partners to improve their relationship collectively and to grow individually.

Acknowledgments

The authors are thankful for the time, support, and access granted by SUSTAINS, cofunded by European Commission and the county council in Uppsala. The paper has benefited from discussions with colleagues in the Development of Online Medical Records and eHealth Services (DOME) consortium. A special thanks to the respondents at Akademiska Sjukhuset in Uppsala.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview template.

[PDF File (Adobe PDF File), 50KB - jmir_v20i1e11_app1.pdf]

References

1. Longtin Y, Sax H, Leape LL, Sheridan SE, Donaldson L, Pittet D. Patient participation: current knowledge and applicability to patient safety. Mayo Clin Proc 2010 Jan;85(1):53-62 [FREE Full text] [doi: 10.4065/mcp.2009.0248] [Medline: 20042562]

2. Barello S, Triberti S, Graffigna G, Libreri C, Serino S, Hibbard J, et al. eHealth for patient engagement: a systematic review. Front Psychol 2016 Jan 08;6:2013 [FREE Full text] [doi: 10.3389/fpsyg.2015.02013] [Medline: 26779108]
3. Ross SE, Lin C. The effects of promoting patient access to medical records: a review. J Am Med Inform Assoc 2003 Mar 01;10(2):129-138. [doi: 10.1197/jamia.M1147]
4. Rexhepi H, Åhlfeldt R, Cajander Å, Huvila I. Cancer patients' attitudes and experiences of online access to their electronic medical records: a qualitative study. Health Informatics J 2016 Jul 19:1-10 Epub ahead of print(forthcoming). [doi: 10.1177/1460458216658778] [Medline: 27440056]
5. 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 Sep;4(9):e006021 [FREE Full text] [doi: 10.1136/bmjopen-2014-006021] [Medline: 25200561]
6. Grünloh C, Cajander Å, Myreteg G. “The Record is Our Work Tool!”—Physicians' framing of a patient portal in Sweden. J Med Internet Res 2016 Jun 27;18(6):e167 [FREE Full text] [doi: 10.2196/jmir.5705] [Medline: 27349531]
7. Clarke V, Braun V. Successful qualitative research: A practical guide for beginners. Los Angeles, London, New Delhi, Singapore, Washington DC: Sage; 2013.
8. Bourgeois FC, Mandl KD, Shaw D, Flemming D, Nigrin DJ. Mychildren’s: Integration of a Personally Controlled Health Record with a Tethered Patient Portal for a Pediatric and Adolescent Population. 2009 Presented at: AMIA Annual Symposium Proceedings; November 14-18, 2009; San Francisco, CA p. 65-69 URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2815447/
9. National High-Level Group for e-Health. 2006. National Strategy for eHealth - Sweden URL: http://workspace.unpan.org/sites/internet/Documents/UNPAN033945.pdf [WebCite Cache ID 6p0aEoBLG]
10. Inera. 2017 Mar 02. Pressmeddelande: 1 miljon användare av e-tjänsten Journalen URL: https://www.inera.se/aktuellt/nyheter/2017/journalen/pressmeddelande-1-miljon-anvandare-av-e-tjansten-journalen/ [accessed 2017-07-13] [WebCite Cache ID 6rvZqDVwS]
11. Veatch RM. Models for ethical medicine in a revolutionary age. What physician-patient roles foster the most ethical relationship? Hastings Cent Rep 1972 Jun;2(3):5-7. [Medline: 4679693]
12. Emanuel EJ, Emanuel LL. Four models of the physician-patient relationship. J Am Med Assoc 1992 Apr 22;267(16):2221-2226. [doi: 10.1001/jama.1992.03480160079038]
13. Deber RB. Physicians in health care management: 7. The patient-physician partnership: changing roles and the desire for information. CMAJ 1994 Jul 15;151(2):171-176 [FREE Full text] [Medline: 8039062]
14. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango). Soc Sci Med 1997;44(5):681-692. [doi: 10.1016/S0277-9536(96)00221-3]
15. Mead N, Bower P. Patient-centredness: a conceptual framework and review of the empirical literature. Soc Sci Med 2000 Oct;51(7):1087-1110. [Medline: 11005395]
16. Szasz TS, Hollender MH. A contribution to the philosophy of medicine: the basic models of the doctor-patient relationship. AMA Arch Intern Med 1956 May;97(5):585-592. [Medline: 13312700]
17. Katz J. The Silent World of Doctor and Patient. Baltimore, MD: Johns Hopkins University Press; 2002.
18. Snow R, Humphrey C, Sandall J. What happens when patients know more than their doctors? Experiences of health interactions after diabetes patient education: a qualitative patient-led study. BMJ Open 2013 Nov 14;3(11):e003583 [FREE Full text] [doi: 10.1136/bmjopen-2013-003583] [Medline: 24231459]
19. 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]
20. Fumagalli LP, Radaelli G, Lettieri E, Bertele' P, Masella C. Patient empowerment and its neighbours: clarifying the boundaries and their mutual relationships. Health Policy 2015 Mar;119(3):384-394. [doi: 10.1016/j.healthpol.2014.10.017] [Medline: 25467286]
21. Deber RB. Physicians in health care management: 8. The patient-physician partnership: decision making, problem solving and the desire to participate. CMAJ 1994 Aug 15;151(4):423-427 [FREE Full text] [Medline: 8055402]
22. Thompson SC, Pitts JS, Schwankovsky L. Preferences for involvement in medical decision-making: situational and demographic influences. Patient Educ Couns 1993 Dec;22(3):133-140. [doi: 10.1016/0738-3991(93)90093-C]
23. Thompson AG. The meaning of patient involvement and participation in health care consultations: a taxonomy. Soc Sci Med 2007 Mar;64(6):1297-1310. [doi: 10.1016/j.socscimed.2006.11.002] [Medline: 17174016]
24. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
25. Williamson L. Patient and citizen participation in health: the need for improved ethical support. Am J Bioeth 2014 May;14(6):4-16 [FREE Full text] [doi: 10.1080/15265161.2014.900139] [Medline: 24809598]
26. Esch T, Mejilla R, Anselmo M, Podtschaske B, Delbanco T, Walker J. Engaging patients through open notes: an evaluation using mixed methods. BMJ Open 2016 Jan 29;6(1):1-11. [doi: 10.1136/bmjopen-2015-010034]
27. Richard AA, Shea K. Delineation of self-care and associated concepts. J Nurs Scholarsh 2011 Sep;43(3):255-264. [doi: 10.1111/j.1547-5069.2011.01404.x] [Medline: 21884371]

28. Omisakin FD, Ncama BP. Self, self-care and self-management concepts: implications for self-management education. Educational Research 2011;2(12):1733-1737 [FREE Full text]
29. Erlingsdottir G, Petersson L. Healthcare personel does not consider the eHealth service “patient online access to their electronic health record” to be beneficial for the work environment and patient safety. 2016 Presented at: NOVO symposium; November 10th-11th, 2016; Reykjavik URL: http://lup.lub.lu.se/record/8dcf3a68-cb1d-4e7c-a93c-d9cc507878fc
30. Aujoulat I, d’Hoore W, Deccache A. Patient empowerment in theory and practice: polysemy or cacophony? Patient Education and Counseling 2007 Apr;66(1):13-20. [doi: 10.1016/j.pec.2006.09.008]

Abbreviations

eHealth: electronic health
EHR: electronic health record
GP: general practitioner
PAEHR: patient accessible electronic health record

Edited by G Eysenbach; submitted 13.07.17; peer-reviewed by T Irizarry, C Pearce, A Blandford; comments to author 11.10.17; revised version received 01.11.17; accepted 02.11.17; published 15.01.18.

Please cite as:
Grünloh C, Myreteg G, Cajander Å, Rexhepi H
“Why Do They Need to Check Me?” Patient Participation Through eHealth and the Doctor-Patient Relationship: Qualitative Study
J Med Internet Res 2018;20(1):e11
URL: http://www.jmir.org/2018/1/e11/
doi:10.2196/jmir.8444
PMID:29335237

Original Paper

The Use of Online Health Forums by Patients With Chronic Cough: Qualitative Study

Ashnish Sinha[1], MBBS, MRes; Tom Porter[1], BA (Hons), MA, PhD; Andrew Wilson[1], MD, FRCP

Norwich Medical School, University of East Anglia, Norwich, United Kingdom

Corresponding Author:
Ashnish Sinha, MBBS, MRes
Norwich Medical School
University of East Anglia
Norwich Research Park
Norwich, NR4 7TJ
United Kingdom
Phone: 44 1603 591257
Email: ashnish@doctors.org.uk

Abstract

Background: Online health discussion forums are used by different patient groups for sharing advice and information. Chronic cough is a common problem, and people with chronic cough use online health forums alongside formal medical therapies.

Objective: The objective of this study was to assess how chronic cough sufferers use online health forums, including the treatment advice they share with one another and the possible clinical uses of online forums in chronic cough.

Methods: Three open-access health forums were searched for threads related to chronic cough. Identified threads were screened against inclusion and exclusion criteria adapted from the British Thoracic Society (BTS) Guidelines related to chronic cough diagnosis. Included data were subjected to qualitative thematic analysis. All study data were cross-validated by a second author and discrepancies were resolved.

Results: In total, 96 threads were included in the analysis, consisting of posts by 223 forum users. Three main themes were identified: the effect of chronic cough on the lives of patients, the treatment advice shared between users, and the provision of support within forums.

Conclusions: Chronic cough symptoms had impacts on multiple aspects of patients' health and well-being. To try and combat these issues, forum users suggested a variety of treatments to one another, ranging from mainstream traditional therapies to odd alternative remedies. The provision of support and empathy were also prominent themes in discussion threads. Online forums themselves may provide increasing benefit to users through the addition of a moderator.

(J Med Internet Res 2018;20(1):e19) doi:10.2196/jmir.7975

KEYWORDS

cough; chronic disease; Internet; eHealth; health impact assessment; information systems; help-seeking behavior; social support

Introduction

The Internet in Health

An ever-increasing number of patients are gaining access to the Internet and harnessing the wealth of information it contains. A 2011 report by the World Health Organization (WHO) estimated that 2 billion people across the world had Internet access, and the current number is likely to be vastly higher [1]. The Internet is a widely available source of health-related information, and [2,3] recent research has also suggested that around 70% of those using the Internet have used it to search for health information, highlighting the high demand for medical information by the public [4].

The Burden of Chronic Cough

People with health complaints routinely use online health forums when seeking information [5]. Chronic cough, defined as a cough which is persistent for more than 8 weeks [6-8], is featured in online health forums. It has an estimated prevalence of 12.7% in Europe and 11.0% in the United States [9]. Yet, the cause of patients' chronic cough often remains undetermined, limiting the scope for clinicians to provide effective treatments [6-8]. Chronic cough has been shown to have an impact on the quality of life (QOL), which has been described in relation to the biopsychosocial model of health [10,11]. Studies have also shown that chronic cough is associated with poorer health-related quality of life (HRQOL) in the domains of social interaction, sleep, and work [12-14], illustrating the potential for chronic cough to disrupt the lives of those whom it affects. Chronic cough also confers burdens beyond the individual, as in the case of parents of children with chronic cough, who report stress and anxiety [15]. The combination of debilitating symptoms and minimal support availability from clinicians often leads patients to seek alternative advice regarding therapies and support, utilizing online forums as a medium for discussion.

Online Health Forums

Online health forums are a source of health information, providing patients with a safe environment to share experiences, seek information, and improve their health knowledge [16-18]. Health forum users have been shown to benefit from online interventions, resulting in greater knowledge about their conditions and greater health activation, with similar efficacy to non-Internet interventions [4,16]. Alongside this, forums provide important opportunities for social support, reassurance, and friendship [18,19]. Recent studies have detailed the content of the information shared in online forums, highlighting its accuracy and identifying the importance of language in promoting discussion [20-23]. The moderation of forums, whereby clinicians or forum staff access the forums and remove harmful posts and avoid repetition, has also been explored [18,24]. Overall, online discussion forums have been shown to be beneficial for users, indicating that further research into their nature and possible utility in medical care is warranted [4,25].

To date, health forum studies have predominantly focused on mental health conditions, but recent research has investigated chronic conditions such as type 2 diabetes mellitus, breast cancer, and stroke [26,27]. Despite its presence in forum discussions, research is yet to explore the use of online health forums in relation to chronic cough. This qualitative study aimed to explore how people living with chronic cough engage with, and make use of, online health forums.

Methods

The study was conducted as a qualitative exploration of three large, open-access online health forums. Forums were identified through a simple search term of “cough health forum” in an online search engine, as in previous studies [19,22]. One author identified threads related to chronic cough within the forums, using a further search term of “chronic cough.” Threads were screened against inclusion and exclusion criteria adapted from the British Thoracic Society (BTS) Guidelines on chronic cough management [8]. Included forum threads were required to be posted in English language and satisfy the criteria for chronic cough. Threads were excluded if patients’ posts suggested an alternative diagnosis or the duration of cough was not specified (Multimedia Appendix 1). This process was continued until data saturation, which was recognized when no new themes were emerging from the coding. We focused particularly on threads discussing chronic cough of unknown etiology, as diagnostic uncertainty is recognized to have an impact on the lived experience of a disease in patients [28,29]. Data were collected between October 2015 and January 2016. The threads themselves ranged in dates from 2002 to 2015 across the three forums, with users still actively participating in these discussions at the time of data collection.

Included forums’ threads were transcribed verbatim into the NVivo 11 program (QSR International Pty Ltd. Version 11, 2015). Threads were analyzed using thematic analysis, with a single author first coding the data into large themes [30]. The data were subsequently coded into subthemes to improve the depth of analysis. The themes and subthemes identified by the primary author were discussed and defended with the other study authors. Then, a secondary author reviewed the threads for the accuracy of coding, but did not recode the data. Discrepancies identified between the study authors were discussed and consensus was reached.

Ethical approval for the study was granted by the host institution’s research ethics committee. Ethical issues were limited by using open-access forums as per previous explorations of Internet research ethics [31-33]. Though previous research has suggested that it is difficult to maintain anonymity in online health forums, the usernames of users posting on the forums were removed, and a unique identifier was given [34]. The authors had no contact with any forum users; thus, no respondent validation could be performed. Risks to forum users from the research were deemed to be minimal.

Results

Figure 1 shows that from the total 223 users, 3 were known medical professionals who were identified because of their usernames and an image from the forum, whereas 1 user in another forum claimed to be a doctor. Moreover, 2 nonmedical moderators were present in forum discussions, and each of them posted in several threads.

Findings

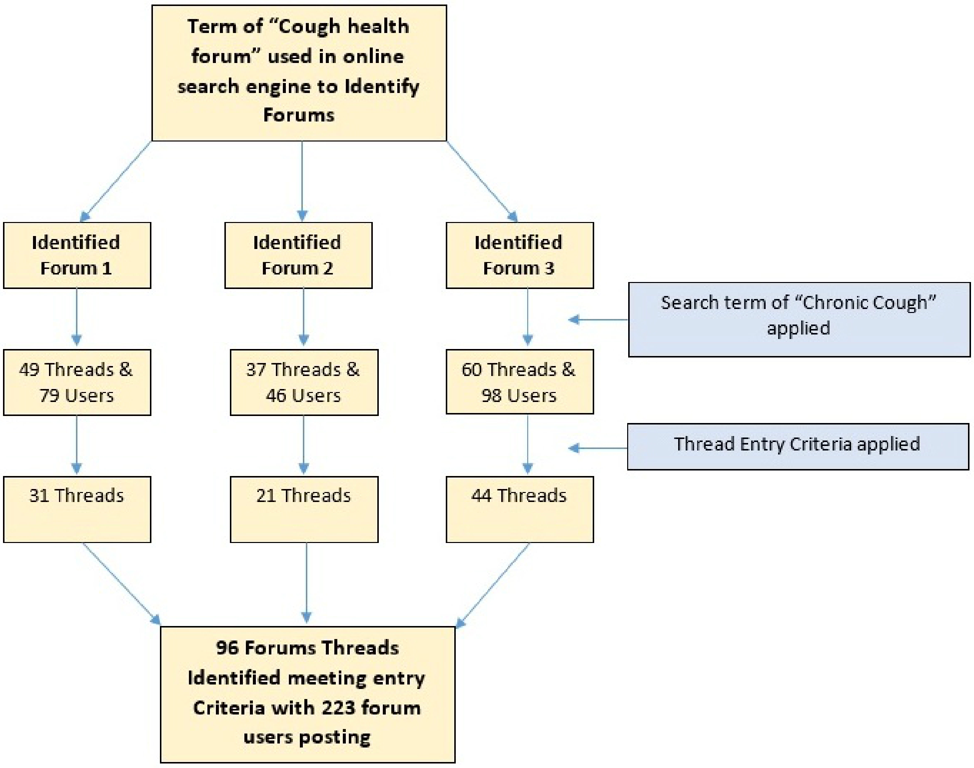
Forum posts consisted of an initial comment or question posed by an original poster (OP), followed by responses and discussion involving responding posters (RP). OP posts not only illustrate the physical impact of chronic cough but also highlight psychological and social consequences. In response to OP posts, RPs offered support in various forms and demonstrated lay knowledge derived from the lived experiences of chronic cough, as well as their appropriation of biomedical knowledge. Forums were a source of distinct forms of support, including suggestions as to traditional and prescribed drug treatments, home remedies, lay referrals, as well as social and emotional support. The following sections outline each of these in turn.

The Impact of Chronic Cough

Physical Consequences

Cough is often seen as an innocuous symptom, yet multiple OPs described adverse physical sequelae resulting from coughing. Of the 96 threads analyzed, 26 contained accounts of physical impairments directly relating to cough.

Figure 1. Flow diagram of the process for thread identification.



The most commonly reported symptoms included bodily aching and vomiting, which were mentioned by 6 and 5 users, respectively. Threads illustrate clearly the distressing and disruptive nature of physical complaints, as User #136 posted: "I am at my wits end," adding "my ribs and my back are very sore from all the coughing."

Another forum user described their spouse's struggle with physical symptoms:

> *...the minute he gets up he coughs until he vomits many times.* [User #133]

Several users detailed yet more debilitating symptoms, including urinary incontinence (2 users) and hoarseness (2 users). Forum users' descriptions of these complaints illustrate just how disruptive the physical consequences of chronic cough may be; one OP said:

> *Last year the cough was so severe I fractured two ribs.* [User #001]

Another user posted that:

> *[He] coughed so hard that I have developed costacondritis [sic] (cracked rib cartilage) twice and ended up in the emergency room, adding that now, I have incontinence [sic] so bad, that I have to urinate every hour.* [User #158]

Psychological Consequences

Adverse physical complaints resulted in psychological distress for many users, with posts detailing how persistent symptoms led to feelings of low mood and frustration. Sleep disturbance, tiredness, and lethargy were also common. One user posted as follows:

> *Now it is a constant cough, is intensified at night (all the night), I can't sleep. Therefore I am very tired, with anxiety, irritable, desperate.* [User #097]

Another user posted succinctly:

> *I feel down, lethargic fatigue* [User #199]

In more extreme instances, forum users said they felt "desperate," and several posters expressed anger toward confidants, with numerous posts detailing frustrating interactions with both family members and health care professionals. One user posted the following:

> *I keep telling my parents to take me to the doctor because I've had this thing for two years, but they always roll their eyes and tell me that I'm making a big deal out of nothing and they think I'm lying. They don't have to live with feeling sick 24/7. They don't know...Someone please tell me what I should do. I really can't take this any longer.* [User #046]

Social Consequences

Several users highlighted the impact of coughing on social relationships and daily activities, with symptoms precluding certain social activities and settings. One user posted the following:

> *I stopped going to church because people look at me like I have something really bad. Sometimes I throw up a lot of mucous and I really cough a lot at night...Sometimes I wet on myself I cough so hard and it makes my head hurt. I can't go anywhere without fear of having these attacks...I've become withdrawn from society.* [User #128]

The following post demonstrates clearly the loss of control and uncertainty delivered by chronic cough: *I can't go anywhere without fear of having these attacks.* In this example, the forum user is concerned about how their symptoms may be perceived by others, which leads them to withdraw from a previously valued social activity. Such social activities and social networks are important domains of QOL, and it is easy to envisage how this forum user's well-being would suffer because of this social withdrawal.

Some OPs were family members posting on behalf of a relative (included parents of children with cough). User #106 was one such parent, whose post detailed experiences of caring for a daughter with cough:

> *My little girl has had an asthma sounding cough for a year on and off. I'm at my wits end as it goes for a week or so comes back for two weeks, In fact I'm pretty sure we have not had a month without coughing, in particular at night. She wakes up coughing at least 3 times at night.*

This forum user continued to explain the impact of their daughter's symptoms, which further emphasize the social consequences of chronic cough:

> *I'm exhausted with it and so is she bless her...I'm on the verge of quitting my job as I can't cope with being up at night and then going to work and I love my job...I feel very down with this and my partner and I argue all the time as we are both tired and things in the house are getting on top of us.* [User #106]

As this excerpt suggests, chronic cough disrupts normal patterns of everyday life, which in turn affect the functioning of social roles and relationships. This forum user explains that her daughter's cough has resulted in strained relations with her partner while also making her own prospects of employment less viable.

Lay Knowledge

In response to original posts, RPs provided various forms of support, including traditional and prescribed drug treatment suggestions, home remedies, and lay referrals. Ensuing threads demonstrate users drawing upon lay knowledge—garnered through the lived experiences of chronic cough and care—to inform, advise, and support.

Biomedical and Prescribed Drug Treatments

In total, forum users recommended 40 differed drug treatments. These recommendations encompassed a variety of treatments ranging from inhalers, such as beta-2 agonists, to antibiotics, and also highly particular treatments for specific conditions, such as Pirfenidone (a treatment for idiopathic pulmonary fibrosis). One user posted the following:

> *It could be rhinitis—allergies. Try taking an antihistamine tablet each night, such as citirizine or loratidine. This worked for me.* [User #036]

This reply is typical in that the forum user draws on their own experience of cough and a successful therapy to make a recommendation to an OP. Moreover, this recommendation is premised upon an inferred diagnosis—rhinitis—which the forum user believes to be an underlying cause of the OP's symptoms. Such assumptions regarding underlying diagnoses were common and often formed the basis for recommended drug treatments. Another example was provided by a user who posted:

> *You may want to see if you respond to an inhaler, and if you do, you definitely have asthma. Asthma has various triggers, so if you can find your triggers, you'll know how to handle the situation.* [User #104]

In other threads, the most common diagnoses offered by forum users were common respiratory conditions, namely, asthma, pneumonia, and pulmonary fibrosis. Appraising the quality of forum users' advice is difficult, but it appears that forum users hold a high level of health knowledge. In most instances, users made sensible suggestions about underlying diagnoses for chronic cough. Furthermore, the advice offered for the amelioration of symptoms was of a good quality, with 33 of 40 recommendations mentioning treatments that were included within the medical guidelines for conditions which commonly cause chronic cough.

Home Remedies

Biomedical treatments comprised the majority of recommendations, yet forum users also suggested home or alternative remedies. The content of home remedies varied widely, but often included specific therapies, such as omega therapy, branded alternative remedies, such as Virastop, and nonprescription tablets or sweets. One user provided a typical example of such recommendations:

> *...so a friend of mine said why dont [sic] you take him to see a omega therapist no medication just healing after 3 sessions we found xx(son) wasnt [sic] coughing as much and after a few months the cough had gone completely (not joking) and the cough never came back, we sat with xx(son) the whole time of his treatment and the lady was just putting her hands on his chest when we asked xx(son) what was happening he said it was like a warm feeling going through his chest, ok we had to pay £20 a session for one hour but it was well worth it.* [User #020]

Other less conventional recommendations included the application of Vicks vapor rub to the soles of the feet. This particular remedy was recommended by 2 users, one of whom posted:

> *I've [sic] recently found out about rubbing a spot of vapour rub/Vicks on the soles of a child's feet at bedtime to stop them coughing. Well I thought I would share this trick as its a miracle and really works. I used it 3 nights in a row on my 6 year old son and he didn't cough all night.* [User #051]

Home remedies were often a point of discussion among forum users who drew on their own experience of cough and of particular remedies to validate or question the efficacy of suggested treatments. Home remedies were generally greeted with enthusiasm by OPs, perhaps due to their unconventional nature and the fact that OPs had often exhausted other conventional means of therapy without resolution and with persisting unexplained symptoms.

Lay Referrals

Many forum users (46) recommended OPs to consult their doctor or health professional rather than attempting to provide treatment suggestions or diagnoses of their own. Such referrals often included suggestions of potential diagnoses or therapies, which were intended to establish a new line of clinical inquiry. One user posted the following:

> *Since you have a teaching job and have spent long hours speaking then it would be advisable to consult an ENT specialist to rule out a vocal cord nodule/growth.* [User #127]

Another user posted the following:

> *Hi User 026, Catch these things early. See your Doctor, slight chance of thrush. We are NOT Doctors. Good luck & Seasons Greetings.* [User #027]

As these examples show, lay referrals often affirmed the expertise of professionals while recognizing the limitations of lay knowledge. Unlike standard clinical interactions, such as doctor-patient consultations, online health forums operate without safeguarding frameworks, meaning that forum users can act and communicate in an entirely unrestricted way. Despite this freedom, lay referrals demonstrate an ethic of self-regulation and an awareness that false or misleading health information may cause harm. Thus, users often stressed the contingent nature of their knowledge and advice, and they regularly referred OPs to the appropriate health professionals.

Social Support

A significant proportion of RP posts involved aspects of social support. Supportive posts were especially common in response to OP posts that displayed distress or which depicted difficult personal or social circumstances.

Supportive posts often invoked shared experiences as a means of displaying empathy and solidarity with OPs. One user posted the following:

> *Hi, sadly it took me six months before I was over it, hope your [sic] soon feeling much better.* [User #012]

Another example was provided by a user who said:

> *Everyone has given you some great advice here. My stress level would be high too if I were going through what you are going through. My heart goes out to you!"* [User #061]

Such posts are intended to give OPs hope that their symptoms would abate or that unexplored therapeutic avenues might resolve their troubles. Supportive posts also illustrate a clear affective connection between RPs and OPs, and support previous research that shows online health forums to be important communities of practice.

Beyond shared experiences, a small number of forum users used humor as a tool to alleviate distress expressed by original posters. One user replied to an original post containing a typo:

> *Secondly if you really are 355 years old it's no wonder you feel bad...hehehe [sic].* [User #200]

Discussion

Study Findings

In this qualitative study of online health forums for chronic cough, we described in detail the multiple adverse physical, psychological, and social effects that were caused due to chronic cough in patients. Subsequently, we explored the treatments and advice suggested by responding forums posters, providing examples of the biomedical and home remedies that users recommended to one another, alongside the idea of lay referral. Finally, we detailed the role of online forums in the provision of support for chronic cough sufferers.

From our analysis, it was evident that chronic cough had a significant impact on the lives of forum users. Its effect was a prominent theme throughout a large number of forum posts. Users highlighted varying levels of physical, psychological, and social impairments, showing that cough symptoms affected all elements of the biopsychosocial model of health; this fits with the previously published research [10,11,14]. Our study supplemented this body of literature by providing an in-depth look at the issues patients were experiencing. Currently, the psychological impacts of cough are more fully documented than the physical consequences. The information we provide details the common physical health consequences of chronic cough, allowing a greater understanding of the symptoms which patients find most concerning. Clinically, this is important as physicians often have few treatments they can provide to combat cough itself. Understanding the subsequent problems provides the opportunity to control these secondary sequelae, improving HRQOL for sufferers. This will also serve to alleviate the frustration that chronic cough patients harbor for medical professionals, who have been unsuccessful in treating their symptoms.

This study also explored the social impacts that cough had on patients and their support network. Users described changing or avoiding their normal activities, as their symptoms induced discrediting stigma. Changes in social activities have previously been reported in relation to chronic cough, whereas stigma has been explored in relation to cystic fibrosis, highlighting cough as a prominent source [11,35]. Our study found similar discussion threads in online health forums with users describing severe limitations to their hobbies or regular activities due to the stigma of their cough symptoms. We also acknowledge the wider impacts of cough, relating to patients' family and friends. This is recognized in other chronic conditions under the family systems theory [36]. There were multiple instances of family members posting concerns for relatives. Similar findings have previously been described [11,14]. Meanwhile, we also found instances of parents posting for their children, demonstrating the parental stress described by Marchant et al [15]. It is important to recognize that the effects of chronic cough extend beyond the sufferer, often impacting their family and support networks. For clinicians, recognizing chronic cough patients' concerns and providing appropriate support will help to alleviate the impact on patients as well as the downstream effects on their family and friends.

Help seeking and the provision of advice were prominent in our data. The seeking of health information through online forums is well documented, and the current evidence suggests that users benefit from this information [4,25]. Accordingly, work on health forum posts has attempted to gauge the accuracy of the information that users provide. Previous evidence suggests that the information provided on health forums is of a good quality, even matching the knowledge of clinicians [22,23]. The analysis in our health forums supported these conclusions. Diagnosis suggestions were common conditions which cause cough, such as asthma or reflux disease. Subsequent treatment suggestions matched these diagnoses, indicating that the information was of good quality. Instances of forum users advising others to reconsult their doctor also highlighted an element of insight and responsibility. It ensured that incorrect information was not provided while also encouraging users with severe symptoms to be assessed by a health professional. With the increasing use of health forums, it is important for clinicians to understand both the good and bad information being shared online. This includes the potential for harmful posts to be left either intentionally through a lack of knowledge from RPs or stemming from an insufficient history being provided by OPs. A proposed method of ensuring forum post quality is the use of medical professional–moderated forums. These have previously been trialed and may allow clinicians to provide accurate information to patients, outside of the normal appointment system [18,37]. Chronic cough is a condition with which patients often visit their doctor multiple times. Our data show forum users avoiding doctors’ appointments after bad experiences or lack of effective treatments, citing them as a waste of time. These patients are lost to follow-up in the medical system, but may frequent online health forums, seeking advice from other sources. Directing them toward a moderated forum, where common questions and concerns are answered by a health professional, could prove time-saving for both the patient and doctor, while also ensuring that they remain engaged with health services and are provided correct, up-to-date information.

In our forums, the suggestion of alternative or so-called home remedies was less prominent than the sharing of traditional treatment advice. Yet, it proved an important facet in forum threads, with multiple users corroborating the suggested treatments. These features have previously been recognized in other studies, which describe the behavior of seeking alternative remedies in online forums, as well as their potential harms [38,39]. In our dataset, the most commonly suggested and corroborated treatments were natural remedies such as herbal oils or suggestions of using alternative medicines. Though alternative remedies are often touted to be effective, it is difficult to justify their use due to uncertainty about their mechanism of action, the extent of the placebo effect, and their potential adverse effects. But clinically, it is important to recognize that patients may try alternative therapies, if traditional treatments have had little efficacy. Our study supplements the current literature by identifying some of the alternative therapies that patients are trying for chronic cough. These are important to identify in case of harmful treatments or potential interactions with prescribed medications. There is also the potential to identify alternative therapies which are effective in alleviating cough symptoms.

The final eminent theme identified was the provision of support through online health forums. Numerous users tagged supportive statements to longer posts or posted purely to convey sympathy or empathy for the OP. Previous research has explored the provision of support in health forums, quantifying the number of supportive posts and their nature [32,40,41]. In our forums, the proportion of supportive posts was outweighed by the number of informative posts fitting the previously reported trends. We also identified the methods used to convey this support. Users often recounted their own similar experiences when conveying empathy, a mechanism identified in other health forum studies [19,32]. Humor was very rarely used, highlighting that online health forums for cough are a platform for serious discussion. In the wider literature, this is reflected by a paucity of evidence for humor in online health forums. Chronic cough sufferers are impacted heavily by their symptoms and require support in dealing with them. Understanding the issues they face allows medical professionals to tailor their services toward these issues, providing more effective support. A moderated forum where these ideals may be performed represents an attractive prospect. It would allow for serious discussion about patients’ symptoms while permitting medical professionals to support patients where possible, and it would also help to prevent irrelevant or incorrect posting which previous authors have shown concern about [38].

Limitations

Our use of open-access, English language, online health forums introduces an element of selection bias due to the prerequisite English language and computer literacy skills required to partake in forum discussions. This also restricts the transferability of our results to subscription-based online health forums. Though the demographics for health forums have been previously reported, we are unable to be sure of how well our forums represent this reported demographic.

Methodologically, we recognize that our search terms for both forum and thread identification alongside our inclusion and exclusion criteria may have resulted in forums or threads related to chronic cough being missed from our analysis. We purposefully excluded chronic related to specific conditions, such as asthma and reflux disease, in our threads, meaning our analysis deals with chronic cough of unknown etiology only. Due to the anonymous nature of posting, we are also unable to corroborate information that is provided, perform triangulation through another data type, or perform respondent validation.

Conclusions

Chronic cough is a widely discussed topic in health forums. Forum users often detail the large impacts that cough has on their daily lives, encompassing all domains of the biopsychosocial model. Clinically, this information may help health professionals to provide therapies for either the cough itself or its sequelae. Information seeking and provision represents key facets of forum use. Cough sufferers often visit forums seeking advice about a possible treatment or diagnosis, with other users providing treatment suggestions or advising reconsultation with a doctor. Moderated forums, where health professionals provide accurate information, are an alternative, proving superior to repeated consultations.

Acknowledgments

The authors would like to acknowledge the University of East Anglia for their support with the project.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Table of inclusion and exclusion criteria as well as a summary of the main themes and subthemes explored in the paper.

[PDF File (Adobe PDF File), 41KB - jmir_v20i1e19_app1.pdf]

References

1. World Health Organization. Switzerland; 2011. Global Observatory for eHealth series - Volume 4 URL: http://www.who.int/goe/publications/ehealth_series_vol4/en/ [accessed 2017-12-29] [WebCite Cache ID 6w4DaWTiL]
2. Powell JA, Darvell M, Gray JA. The doctor, the patient and the world-wide web: how the internet is changing healthcare. J R Soc Med 2003 Feb;96(2):74-76 [FREE Full text] [Medline: 12562977]
3. Jadad AR, Gagliardi A. Rating health information on the Internet: navigating to knowledge or to Babel? J Am Med Assoc 1998 Feb 25;279(8):611-614. [Medline: 9486757]
4. Grosberg D, Grinvald H, Reuveni H, Magnezi R. Frequent surfing on social health networks is associated with increased knowledge and patient health activation. J Med Internet Res 2016 Aug 10;18(8):e212 [FREE Full text] [doi: 10.2196/jmir.5832] [Medline: 27511272]
5. Sudau F, Friede T, Grabowski J, Koschack J, Makedonski P, Himmel W. Sources of information and behavioral patterns in online health forums: observational study. J Med Internet Res 2014 Jan 14;16(1):e10 [FREE Full text] [doi: 10.2196/jmir.2875] [Medline: 24425598]
6. Gibson P, Wang G, McGarvey L, Vertigan AE, Altman KW, Birring SS, CHEST Expert Cough Panel. Treatment of Unexplained Chronic Cough: CHEST Guideline and Expert Panel Report. Chest 2016 Jan;149(1):27-44. [doi: 10.1378/chest.15-1496] [Medline: 26426314]
7. Brown KK. Chronic cough due to chronic interstitial pulmonary diseases: ACCP evidence-based clinical practice guidelines. Chest 2006 Jan;129(1 Suppl):180S-185S. [doi: 10.1378/chest.129.1_suppl.180S] [Medline: 16428708]
8. Morice AH, McGarvey L, Pavord I, British Thoracic Society Cough Guideline Group. Recommendations for the management of cough in adults. Thorax 2006 Sep;61 Suppl 1:i1-24 [FREE Full text] [doi: 10.1136/thx.2006.065144] [Medline: 16936230]
9. Song W, Chang YS, Faruqi S, Kim JY, Kang MG, Kim S, et al. The global epidemiology of chronic cough in adults: a systematic review and meta-analysis. Eur Respir J 2015 May;45(5):1479-1481 [FREE Full text] [doi: 10.1183/09031936.00218714] [Medline: 25657027]
10. Dicpinigaitis PV, Tso R, Banauch G. Prevalence of depressive symptoms among patients with chronic cough. Chest 2006 Dec;130(6):1839-1843. [doi: 10.1378/chest.130.6.1839] [Medline: 17167006]
11. Everett CF, Kastelik JA, Thompson RH, Morice AH. Chronic persistent cough in the community: a questionnaire survey. Cough 2007 Mar 23;3:5 [FREE Full text] [doi: 10.1186/1745-9974-3-5] [Medline: 17381836]
12. Irwin RS, French CT, Fletcher KE. Quality of life in coughers. Pulm Pharmacol Ther 2002;15(3):283-286. [doi: 10.1006/pupt.2002.0347] [Medline: 12099779]
13. Ma W, Yu L, Wang Y, Li X, Lü H, Qiu Z. Changes in health-related quality of life and clinical implications in Chinese patients with chronic cough. Cough 2009 Sep 25;5:7 [FREE Full text] [doi: 10.1186/1745-9974-5-7] [Medline: 19781068]
14. Chamberlain SA, Garrod R, Douiri A, Masefield S, Powell P, Bücher C, et al. The impact of chronic cough: a cross-sectional European survey. Lung 2015 Jun;193(3):401-408. [doi: 10.1007/s00408-015-9701-2] [Medline: 25787221]
15. Marchant JM, Newcombe PA, Juniper EF, Sheffield JK, Stathis SL, Chang AB. What is the burden of chronic cough for families? Chest 2008 Aug;134(2):303-309. [doi: 10.1378/chest.07-2236] [Medline: 18641100]
16. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014 Feb 14;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]
17. Barney LJ, Griffiths KM, Banfield MA. Explicit and implicit information needs of people with depression: a qualitative investigation of problems reported on an online depression support forum. BMC Psychiatry 2011 May 18;11:88 [FREE Full text] [doi: 10.1186/1471-244X-11-88] [Medline: 21592335]
18. Kendal S, Kirk S, Elvey R, Catchpole R, Pryjmachuk S. How a moderated online discussion forum facilitates support for young people with eating disorders. Health Expect 2017 Feb;20(1):98-111 [FREE Full text] [doi: 10.1111/hex.12439] [Medline: 26725547]

19. Bhamrah G, Ahmad S, NiMhurchadha S. Internet discussion forums, an information and support resource for orthognathic patients. Am J Orthod Dentofacial Orthop 2015 Jan;147(1):89-96. [doi: 10.1016/j.ajodo.2014.08.020] [Medline: 25533076]
20. Durant KT, McCray AT, Safran C. Identifying temporal changes and topics that promote growth within online communities: a prospective study of six online cancer forums. Int J Comput Models Algorithms Med 2011 Apr;2(2):1-22 [FREE Full text] [doi: 10.4018/jcmam.2011040101] [Medline: 22022636]
21. Crook B, Glowacki EM, Love B, Jones BL, Macpherson CF, Johnson RH. Hanging by a thread: exploring the features of nonresponse in an online young adult cancer survivorship support community. J Cancer Surviv 2016 Feb;10(1):185-193. [doi: 10.1007/s11764-015-0465-8] [Medline: 26123806]
22. Cole J, Watkins C, Kleine D. Health advice from internet discussion forums: how bad is dangerous? J Med Internet Res 2016 Jan 6;18(1):e4 [FREE Full text] [doi: 10.2196/jmir.5051] [Medline: 26740148]
23. Nelson S, Hwang KO, Bernstam EV. Comparing clinician knowledge and online information regarding Alli (Orlistat). Int J Med Inform 2009 Nov;78(11):772-777 [FREE Full text] [doi: 10.1016/j.ijmedinf.2009.07.003] [Medline: 19716762]
24. Wise K, Hamman B, Thorson K. Moderation, response rate, and message interactivity: features of online communities and their effects on intent to participate. J Comp Mediated Comm 2006 Oct;12(1):24-41. [doi: 10.1111/j.1083-6101.2006.00313.x]
25. Green RA, Farchione D, Hughes DL, Chan SP. Participation in asynchronous online discussion forums does improve student learning of gross anatomy. Anat Sci Educ 2014;7(1):71-76. [doi: 10.1002/ase.1376] [Medline: 23733524]
26. Chen AT. Exploring online support spaces: using cluster analysis to examine breast cancer, diabetes and fibromyalgia support groups. Patient Educ Couns 2012 May;87(2):250-257. [doi: 10.1016/j.pec.2011.08.017] [Medline: 21930359]
27. De Simoni A, Shanks A, Balasooriya-Smeekens C, Mant J. Stroke survivors and their families receive information and support on an individual basis from an online forum: descriptive analysis of a population of 2348 patients and qualitative study of a sample of participants. BMJ Open 2016 Apr 6;6(4):e010501 [FREE Full text] [doi: 10.1136/bmjopen-2015-010501] [Medline: 27053271]
28. Nettleton S, Watt I, O'Malley L, Duffey P. Understanding the narratives of people who live with medically unexplained illness. Patient Educ Couns 2005 Feb;56(2):205-210. [doi: 10.1016/j.pec.2004.02.010] [Medline: 15653250]
29. Nettleton S. 'I just want permission to be ill': towards a sociology of medically unexplained symptoms. Soc Sci Med 2006 Mar;62(5):1167-1178. [doi: 10.1016/j.socscimed.2005.07.030] [Medline: 16135395]
30. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
31. Convery I, Cox D. A review of research ethics in internet-based research. Pract Res High Educ 2012;6(1):50-57 [FREE Full text]
32. Kaufman S, Whitehead KA. Producing, ratifying, and resisting support in an online support forum. Health (London) 2016 Feb 5:-. [doi: 10.1177/1363459315628043] [Medline: 26851264]
33. Eysenbach G, Till JE. Ethical issues in qualitative research on internet communities. BMJ 2001;323:1103. [doi: 10.1136/bmj.323.7321.1103]
34. Bobicev V, Sokolova M, El Emam K, Jafer Y, Dewar B, Jonker E, et al. Can anonymous posters on medical forums be reidentified? J Med Internet Res 2013 Oct 3;15(10):e215 [FREE Full text] [doi: 10.2196/jmir.2514] [Medline: 24091380]
35. Pakhale S, Armstrong M, Holly C, Edjoc R, Gaudet E, Aaron S, et al. Assessment of stigma in patients with cystic fibrosis. BMC Pulm Med 2014 May 1;14:76 [FREE Full text] [doi: 10.1186/1471-2466-14-76] [Medline: 24884812]
36. Kerr ME. One Family's Story: A Primer on Bowen Theory. Washington, DC: The Bowen Center for the Study of the Family; 2000.
37. Lindsay S, Smith S, Bellaby P, Baker R. The health impact of an online heart disease support group: a comparison of moderated versus unmoderated support. Health Educ Res 2009 Aug;24(4):646-654. [doi: 10.1093/her/cyp001] [Medline: 19251770]
38. McGregor F, Somner JE, Bourne RR, Munn-Giddings C, Shah P, Cross V. Social media use by patients with glaucoma: what can we learn? Ophthalmic Physiol Opt 2014 Jan;34(1):46-52. [doi: 10.1111/opo.12093] [Medline: 24325434]
39. Tozzi AE, Mingarelli R, Agricola E, Gonfiantini M, Pandolfi E, Carloni E, et al. The internet user profile of Italian families of patients with rare diseases: a web survey. Orphanet J Rare Dis 2013 May 16;8:76 [FREE Full text] [doi: 10.1186/1750-1172-8-76] [Medline: 23680013]
40. Deetjen U, Powell JA. Informational and emotional elements in online support groups: a Bayesian approach to large-scale content analysis. J Am Med Inform Assoc 2016 May;23(3):508-513. [doi: 10.1093/jamia/ocv190] [Medline: 26911815]
41. Moore D, Ayers S. Virtual voices: social support and stigma in postnatal mental illness Internet forums. Psychol Health Med 2017 Jun;22(5):546-551. [doi: 10.1080/13548506.2016.1189580] [Medline: 27218265]

Abbreviations

BTS: British Thoracic Society
HRQOL: health-related quality of life
OP: original poster
QOL: quality of life

RP: responding poster
UEA: University of East Anglia
WHO: World Health Organization

Edited by G Eysenbach; submitted 03.05.17; peer-reviewed by L Sbaffi, J Hirschey; comments to author 22.06.17; revised version received 29.07.17; accepted 08.09.17; published 24.01.18.

Please cite as:
Sinha A, Porter T, Wilson A
The Use of Online Health Forums by Patients With Chronic Cough: Qualitative Study
J Med Internet Res 2018;20(1):e19
URL: http://www.jmir.org/2018/1/e19/
doi:10.2196/jmir.7975
PMID:29367181

Original Paper

Prediction of Incident Hypertension Within the Next Year: Prospective Study Using Statewide Electronic Health Records and Machine Learning

Chengyin Ye[1,2*], PhD; Tianyun Fu[3*], BS; Shiying Hao[4,5*], PhD; Yan Zhang[6], MD; Oliver Wang[3], BA; Bo Jin[3], MS; Minjie Xia[3], BS; Modi Liu[3], MS; Xin Zhou[7], MD; Qian Wu[8], BS; Yanting Guo[2,9], BS; Chunqing Zhu[3], MS; Yu-Ming Li[7], MD; Devore S Culver[10], MM; Shaun T Alfreds[10], MBA; Frank Stearns[3], MHA; Karl G Sylvester[2], MD; Eric Widen[3], MHA; Doff McElhinney[4,5*], MD; Xuefeng Ling[2,5,11*], PhD

[1]Department of Health Management, Hangzhou Normal University, Hangzhou, China

[2]Department of Surgery, Stanford University, Stanford, CA, United States

[3]HBI Solutions Inc, Palo Alto, CA, United States

[4]Department of Cardiothoracic Surgery, Stanford University, Stanford, CA, United States

[5]Clinical and Translational Research Program, Betty Irene Moore Children's Heart Center, Lucile Packard Children's Hospital, Stanford, CA, United States

[6]Department of Oncology, The First Hospital of Shijiazhuang, Shijiazhuang, China

[7]Tianjin Key Laboratory of Cardiovascular Remodeling and Target Organ Injury, Pingjin Hospital Heart Center, Tianjin, China

[8]China Electric Power Research Institute, Beijing, China

[9]School of Management, Zhejiang University, Hangzhou, China

[10]HealthInfoNet, Portland, ME, United States

[11]Health Care Big Data Center, School of Public Health, Zhejiang University, Hangzhou, China

[*]these authors contributed equally

Corresponding Author:
Xuefeng Ling, PhD
Department of Surgery
Stanford University
S370, 300 Pasteur Drive
Stanford, CA, 94305
United States
Phone: 1 6504279198
Email: bxling@stanford.edu

Related Article:

Comment in: https://www.jmir.org/2021/2/e10969

Abstract

Background: As a high-prevalence health condition, hypertension is clinically costly, difficult to manage, and often leads to severe and life-threatening diseases such as cardiovascular disease (CVD) and stroke.

Objective: The aim of this study was to develop and validate prospectively a risk prediction model of incident essential hypertension within the following year.

Methods: Data from individual patient electronic health records (EHRs) were extracted from the Maine Health Information Exchange network. Retrospective (N=823,627, calendar year 2013) and prospective (N=680,810, calendar year 2014) cohorts were formed. A machine learning algorithm, XGBoost, was adopted in the process of feature selection and model building. It generated an ensemble of classification trees and assigned a final predictive risk score to each individual.

Results: The 1-year incident hypertension risk model attained areas under the curve (AUCs) of 0.917 and 0.870 in the retrospective and prospective cohorts, respectively. Risk scores were calculated and stratified into five risk categories, with 4526 out of 381,544 patients (1.19%) in the lowest risk category (score 0-0.05) and 21,050 out of 41,329 patients (50.93%) in the highest risk category (score 0.4-1) receiving a diagnosis of incident hypertension in the following 1 year. Type 2 diabetes, lipid disorders, CVDs, mental illness, clinical utilization indicators, and socioeconomic determinants were recognized as driving or associated features of incident essential hypertension. The *very high risk* population mainly comprised elderly (age>50 years) individuals with multiple chronic conditions, especially those receiving medications for mental disorders. Disparities were also found in social determinants, including some community-level factors associated with higher risk and others that were protective against hypertension.

Conclusions: With statewide EHR datasets, our study prospectively validated an accurate 1-year risk prediction model for incident essential hypertension. Our real-time predictive analytic model has been deployed in the state of Maine, providing implications in interventions for hypertension and related diseases and hopefully enhancing hypertension care.

(J Med Internet Res 2018;20(1):e22) doi:10.2196/jmir.9268

KEYWORDS

hypertension; risk assessment; electronic health records; multiple chronic conditions; mental disorders; social determinants of health

Introduction

Background

Hypertension affects more than 85 million American adults and over 1.1 billion people worldwide, and elevated blood pressure is one of the leading modifiable metabolic risk factors for noncommunicable diseases and global mortality [1,2]. For example, in some age groups, the risk of cardiovascular disease (CVD) can double with each 20-10 mmHg increase of blood pressure [3]. Therefore, efficient management of blood pressure is a critical step toward reducing morbidity and mortality from chronic diseases such as coronary heart disease and stroke. Moreover, inadequate management of hypertension, including failure to diagnosis, treat, and control high blood pressure, contributes enormous medical, socioeconomic, and human costs [4]. In the United States, the average annual total cost associated with high blood pressure was US $51.2 from 2012 to 2013, with US $37.2 billion in direct costs from health care services and medications and US $3.9 billion in indirect costs owing to lost productivity from morbidity and mortality [1].

Commonly identified risk factors for progression to hypertension include age, gender, body mass index, obesity, stress, lipoproteins, cholesterol, physical inactivity, smoking, and family history [5]. Evidence from clinical trials suggests that early prevention of hypertension with lifestyle modification or drug treatment among individuals without hypertension but having risk factors or prehypertension may reduce the incidence of hypertension in the short-term and for several years after [6-8]. On the other hand, a potentially desirable and operable alternative to universal implementation of such measures is to stratify individuals into different risk groups by applying prediction tools, then target the highest risk segment of the population for subsequent lifestyle intervention or medical treatment, and eventually prevent their progression from high-risk status to actual hypertension [9].

Possible Limitations of Existing Hypertension Prediction Models

In the last two decades, more than 15 high-quality hypertension risk prediction models were constructed and compared in terms of study design, model performance, calibration ability, as well as impact on decision making and outcomes of care [10]. These models all achieved acceptable to good discrimination with an area under the curve (AUC) between 0.71 and 0.81. However, as they were formed using traditional statistical methods (eg, Cox regression or logistic regression), most existing hypertension prediction models are limited by either small sample size (ie, using data from only a single medical facility), relatively few (<15) traditional risk factors, or lack of ability to monitor real-time changes in predictors and therefore hypertension risk [11].

On the other hand, the use of electronic health records (EHRs) is becoming increasingly common in hospitals and clinics [12]. By tracking almost all aspects of a patient's care over time, EHR data contain an individual's integrated and comprehensive clinical history and can be used to build risk models that can potentially help predict disease progression, revealing the evolution of disease, and thus, offering great promise for accelerating clinical research and predictive analysis on a population [13,14]. As the versatility of EHR-based datasets provides more generalized prediction results with high levels of confidence, the clinical construction of EHR-based risk prediction models has become more effective and is in higher demand [15]. Although several EHR-based prediction models have been successfully built for other diseases (eg, type 2 diabetes, kidney disease, and cancer) [14,16-26], such models are still relatively unexplored for the prediction of hypertension.

Aim of This Study

Given that the EHR data have many dimensions and are usually sparse and subject to random errors, machine learning approaches are recognized as good choices for constructing EHR-based risk models based on their ability to select impactful predictors automatically from hundreds of features and their robustness to feature correlation and random errors [27-29]. Therefore, to take advantage of both EHR data and machine learning approaches, our study aimed to construct and prospectively validate a new hypertension risk prediction model that is able to utilize an individual's current 1-year clinical information to capture previously ignored but potentially powerful predictors directly from EHRs, including patients' current health conditions, chronic disease and medication history, clinical utilization measures, and social determinants.

It is hoped that this new approach could ultimately predict the probability of receiving a new hypertension diagnosis in the near term (ie, during the next 1 year) with improved accuracy.

Methods

The Health Information Exchange Dataset of Maine

The analyzed dataset was derived from EHRs of all patients that visited any of 35 hospitals, 34 federally qualified health centers, and more than 400 ambulatory practices in the state of Maine from January 1, 2013 to December 31, 2015. This dataset covered almost 95% of the population of Maine and is a subset of the health information exchange (HIE) network provided by the HealthInfoNet organization. This study was approved by the institutional review board of Stanford University, and all personal privacy information was well protected and removed during the process of analysis and publication. Samples were excluded if they had any diagnosis of secondary or gestational hypertension, leaving a total of 1.5 million individuals in the dataset for the analysis. The detailed inclusion and exclusion criteria were demonstrated in the study design workflow (Figure 1).

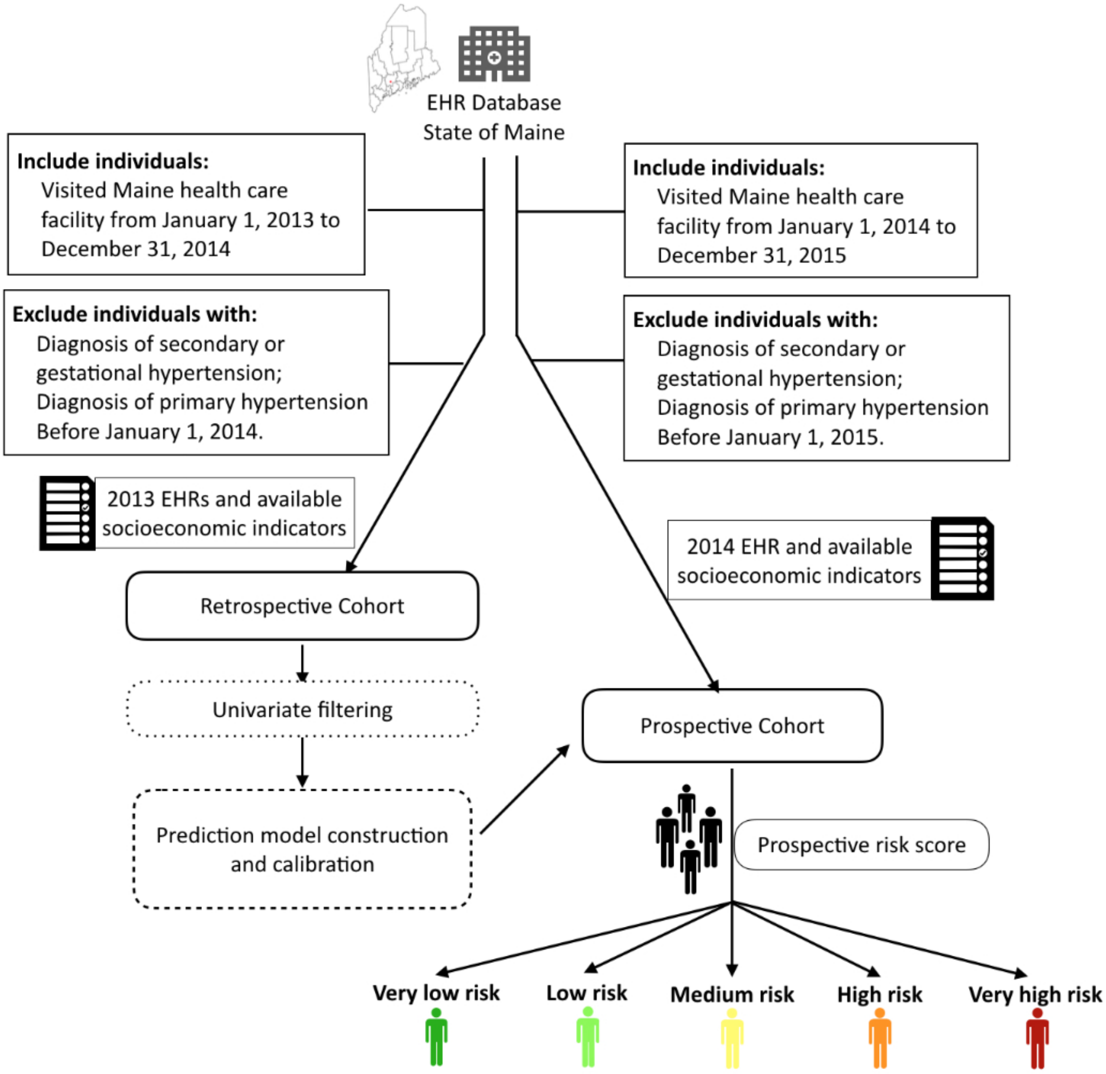
Definition of Hypertension

Hypertension in this study was defined by the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) diagnosis codes from category 401, which refers specifically to essential (or primary) hypertension. For the prediction modeling cohort, cases in the retrospective population referred to patients who received a new diagnosis of essential hypertension during the calendar year of 2014 (from January 1, 2014 to December 31, 2014), whereas cases in the prospective cohort were patients receiving an incident diagnosis of essential hypertension during the 2015 calendar year (from January 1, 2015 to December 31, 2015). Along with the implementation of ICD-10-CM since October 1, 2015, an ICD-10-CM category of I10 was used to define the diagnosis of primary hypertension, which is equivalent to the ICD-9-CM category 401 according to the ICD-10-CM General Equivalence Mapping (GEM), a crosswalk between the two code standards maintained by the Center for Medicare Services and the Centers for Disease Control and Prevention [30].

Cohorts

Given that our prediction model was constructed to predict an individual's hypertension risk during the following 1 year based on his or her medical records from the current year, feature profiles for the retrospective cohort were extracted from the clinical and health historical record of 2013, giving a total of 823,627 patients, 92,512 of whom developed hypertension in the year 2014.

Figure 1. Workflow diagram depicting model construction and evaluation. The retrospective cohort consisted of 823,627 individuals with electronic health record (EHR) profiles extracted from 2013, 92,512 of whom (cases) developed hypertension in 2014. The validation cohort consisted of 680,810 individuals, with EHR profiles extracted from 2014, 60,065 of whom received a new hypertension diagnosis in 2015.



Samples were excluded from the retrospective cohort if there was any record of a hypertension diagnosis before January 1, 2014. Similarly, the prospective cohort was extracted from the clinical and health historical record of 2014, offering a total of 680,810 samples and 60,065 cases with incident hypertension in the year 2015. Samples were excluded in prospective cohort if there was any record of a hypertension diagnosis before January 1, 2015.

Prediction Model Construction and Evaluation

Features Preprocessing and Selection

Various categories of data were extracted from the original health records, including demographics, laboratory and radiographic test results, essential and secondary diagnoses and procedures, outpatient medication prescriptions, clinical utility records, as well as a number of accessible socioeconomic variables extracted from the US census and United States Department of Agriculture (USDA) websites [31,32]. Overall, more than 15,000 features were recruited in our original data pool. Among them, laboratory and radiographic test results were coded by logical observation identifier names and codes, outpatient medication prescriptions were coded by the National Drug Code, and primary and secondary diagnoses and procedures were coded by ICD-9-CM. All ICD-10-CM codes were mapped to ICD-9-CM using the GEM tool to keep the data consistent. Some socioeconomic features are directly downloaded from the US census website using the advanced search tool under the *American Fact Finder* category, and others were derived from the USDA website using the Web-based mapping tool, *Food Environment Atlas* (see Multimedia Appendix 1). Using zone improvement plan (ZIP) code or county or tract, all these socioeconomic variables are integrated into our knowledge base.

Before feature selection, the k-nearest neighbors (KNN) approach was used to impute missing data. For a given patient with missing values, the KNN method identified the k-nearest patients based on Euclidean distance. Using these patients, missing values were then replaced using a majority vote for discrete variables and weighted means for continuous features. One advantage of using this method is that missing values in all features are imputed simultaneously without the need to treat features individually [33].

Before the machine-learning algorithm, a univariate correlation filtering step was introduced to remove all features that are not directly related to the outcome variable [34]. In general, the Cochran-Mantel-Haenszel test is applied for binary features by investigating the association between the feature and outcome under age-group strata using the following formula: $R=\sum^{K}_{i=1}A_iD_i/T_i/\ \sum^{K}_{i=1}B_iC_i/T_i$, where A_i, B_i, C_i, and D_i are the counts of individuals in 2x2 contingency table of the *i*- th age group. In addition, the Cochran-Armitage trend test is used to test ordinal variables, and univariate logistic regression is used for continuous variables. After applying the criteria of $P \leq .05$, 798 out of 15,280 features survived from this screening procedure and were recruited in the following stages of XGBoost algorithm.

Model Construction

The retrospective cohort was utilized for construction of the prediction model. This process was accomplished in two phases: (1) a data training algorithm to generate predictive estimates and (2) a calibration stage to produce risk scores for each patient.

For the first stage of data training, a supervised machine learning and data mining tool, XGBoost [35], was applied to develop the prediction model based on features survived in the univariate phase. Generally speaking, this algorithm generates an ensemble of classification trees and sums the scores in the corresponding leaves of each tree to calculate a final predictive estimate $y\hat{}_i$ for the *i*- th instance ($i=1,..., n$), as demonstrated in equation 1, where each f_k corresponds to an independent classification tree:

$$y\hat{}_i = ø(x_i) = \sum^{K}_{k=1} f_k(x_i) \quad (1)$$

For this study, the depth of each tree was set to be 5 and K equaled 500. The model at the *t*- th iteration was trained to minimize the following objective, where l is a differentiable convex loss function that not only measures the difference between the target y_i and the prediction $y\hat{}_i^{(t-1)}$ of the *i*-th instance at the *t-1*- th iteration but also takes the f_t that improves the model most into account:

$$L^{(t)} = \sum^{n}_{i=1} l(\,y_i, y\hat{}_i^{(t-1)} + f_t(x_i)) + \Omega\ (f_t) \quad (2)$$

The term Ω penalizes the complexity of the regression tree functions to avoid issues of overfitting. The approximate greedy algorithm was adopted to split trees and sort and pick features on each node according to percentiles of distribution. Splitting points were selected to optimize purity at the next splitting level.

For the second phase of calibration, $y\hat{}$ estimates were mapped to positive predictive values (PPVs) in the retrospective cohort. A PPV of a corresponding predictive estimate $y\hat{}$ was defined as the proportion of incident hypertension events in the cohort having predictive estimates the same as or larger than this $y\hat{}$. Thus, PPVs could also be interpreted as risk scores that measured probability of receiving a new diagnosis of essential hypertension within the next 1 year among individuals having predictive estimates the same as or larger than this $y\hat{}$.

Model Evaluation

In the process of model evaluation, hypertension risk score expressed as PPVs were calculated and assigned to each individual in the prospective cohort. Following that, we further ranked individuals by their risk scores from low to high and allocated them into five distinct risk categories, manifesting individuals' chance of developing essential hypertension within the next 1 year to be either *very low*, *low*, *medium*, *high*, or *very high*. Performance of the proposed 1-year hypertension risk prediction model was investigated carefully within each risk category in terms of sensitivity, specificity, and PPV. The receiver operating characteristic curve and the validated AUC value were derived to evaluate the performance of the derived prediction model.

For these highly weighted features, age-sex adjusted odds ratios (ORs) between cases and controls were calculated along with 95% CI using logistic regression. Distribution patterns of impactful risk predictors were explored and compared among different risk categories, uncovering predictors' characteristics under our constructed prediction model. Moreover, survival analyses such as multivariable Cox regression were employed for the purpose of subpopulation comparison. Spearman rank correlations were performed to assess the correlation between socioeconomic features and the next 1-year risk of hypertension.

Results

Demographic Baseline

Baseline demographic features and hypertension-related disease conditions in retrospective and prospective cohort, summarized in Table 1, indicate that most characteristics were similarly distributed between the two cohorts. When considering hypertension relevant diagnostic diseases, CVDs, disorders of lipid metabolism, type 2 diabetes, and chronic obstructive pulmonary disease (COPD) were the most common diseases in the nonhypertensive retrospective and prospective cohorts. Furthermore, the distributions of other chronic diseases such as anemia, idiopathic hypersomnia, prehypertension, and chronic kidney disease (CKD) were balanced between the retrospective and prospective populations.

Table 1. Baseline characteristics of the retrospective and prospective cohorts.

Characteristic	Training cohort (N=823,627) n (%)	Validation cohort (N=680,810) n (%)
Age (years)		
<35	389,081 (47.23)	403,932 (59.30)
35-49	202,906 (24.64)	158,580 (23.29)
50-64	118,624 (14.40)	71,687 (10.53)
≥65	112,709 (13.68)	46,611 (6.85)
Gender		
Male	366,859 (44.54)	294,430 (43.25)
Female	456,768 (55.46)	386,380 (56.75)
Race		
White	781,457 (94.88)	634,107 (93.14)
Black	9060 (1.11)	10,212 (1.50)
Asian	8978 (1.09)	6332 (0.93)
Other	24,050 (2.92)	30,159 (4.43)
Diagnostic disease		
Cardiovascular diseases	50,769 (6.16)	26,259 (3.86)
Disorders of lipid metabolism	22,223 (2.70)	22,063 (3.24)
Type 2 diabetes	23,464 (2.85)	17,999 (2.64)
COPD[a]	8391 (1.02)	8534 (1.25)
Acquired hemolytic anemia	6508 (0.79)	6289 (0.92)
Liver disorders	6194 (0.75)	4121 (0.61)
Idiopathic hypersomnia	5141 (0.62)	2991 (0.44)
Prehypertension	2439 (0.30)	1877 (0.28)
Chronic nephritis	3055 (0.37)	1713 (0.25)
Chronic kidney disease	2389 (0.29)	1404 (0.21)
Hypopotassemia	1245 (0.15)	765 (0.11)
Hyposmolality or hyponatremia	549 (0.07)	381 (0.06)
Diagnosed as primary hypertension after 1 year	92,512 (11.23)	60,065 (8.82)

[a] COPD: chronic obstructive pulmonary disease.

Model Performance

By adopting the machine learning tool XGBoost on the EHR-based data, our prediction model reached a fitted AUC of 0.917 in the retrospective cohort and a predicted AUC of 0.870 in the independent prospective cohort (Figure 2). A total of 381,544 individuals were labeled as *very low risk* (ie, received risk score 0-0.05), and <1.19% (4526/381,544) were affected by hypertension in the next 1 year. In contrast, for 41,329 individuals identified in the *very high risk* category (ie, risk score >0.4), more than 50.93% (21,050/41,329) received a confirmatory diagnosis of essential hypertension during the next 1 year. Among 60,065 patients with confirmed hypertension in the next 1 year, more than one-third (35.04%, 21,050/60,065) were correctly classified into the *very high risk* category, and only 7.54% (4526/60,065) of them were falsely assigned to the *very low risk* population (see Multimedia Appendix 2).

To further evaluate the model’s discriminative ability based on the five risk categories, we estimated each category’s time-to-hypertension curve using univariable Cox regression. As a result, five distinct survival curves were created and well stratified in terms of hypertension hazard (*P*<.001), giving a hazard ratio (HR) for the *very high risk* category as high as 60.8 (95% CI 58.8-62.8) relative to the *very low risk* group (Figure 2).

With the original pool of 15,280 features merging from EHRs and socioeconomic indicators, 798 features passed the univariate selection phase and were recruited as candidates of predictors for the following machine learning algorithm. During the process of model construction, XGBoost adopted the approximate greedy algorithm to split trees by sorting and picking features on each node to optimize purity at each splitting level. Finally, it automatically recruited a total of 169 features as predictors into the prediction model, consisting of two demographic features, 14 socioeconomic characteristics, 30 diagnostic diseases, six laboratory tests, 98 medication prescriptions, and 19 clinical utilization measures (Figure 3). The most impactful 80 features are listed in Multimedia Appendix 3, as well as their age-sex adjusted ORs or coefficients and 95% CIs derived from the prospective cohort. Besides age and sex, the meaningful features mainly comprised hypertension-related disease diagnoses, medications for these related diseases, clinical utilization measurements, and socioeconomic indicators. For diagnosed diseases, prehypertension, type 2 diabetes, CVDs (combination of congestive heart failure, atherosclerotic heart disease, coronary heart disease, and myocardial infarction), and idiopathic hypersomnia were recognized as the conditions most associated with hypertension, all with an OR>3.00 in prospective cohort. In terms of medication prescriptions, other than those for CVDs, type 2 diabetes, and disorders of lipid metabolism, our prediction model also captured consumed medications for treatment of mental health conditions, including drugs for depression, anxiety, and schizophrenia. From the clinical utilization domain, inpatient admissions and outpatient visits during last year and total estimated health care costs for the patient over last year were significant. Furthermore, social determinants such as education level, type of health insurance, and environmental factors related to dietary habits and physical activities were also detected by the model as powerful predictors of incident hypertension within the next 1 year.

Significant Features

To further validate the impact of recognized risk factors on the risk of essential hypertension, we carefully investigated the *very high risk* population in the prospective cohort and compared it with the *very low risk* category, revealing the unique characteristics of *very high risk* individuals.

Demographic Features

Age and gender were recognized as the most impactful demographic features in our hypertension risk prediction model. To follow the regulations of Health Insurance Portability and Accountability Act, all ages over 89 years were combined into one category to protect patients' health information.

Figure 2. (a) The receiver operating characteristic (ROC) curve derived from the prospective cohort (left). (b) Survival curves (right) of the five risk categories identified by the 1-year hypertension risk prediction model. Five risk categories were defined: very low (risk score<0.05), low (risk score 0.05-0.1), medium (risk score 0.1-0.2), high (risk score 0.2-0.4), and very high (risk score >0.4). HR: hazard ratio.

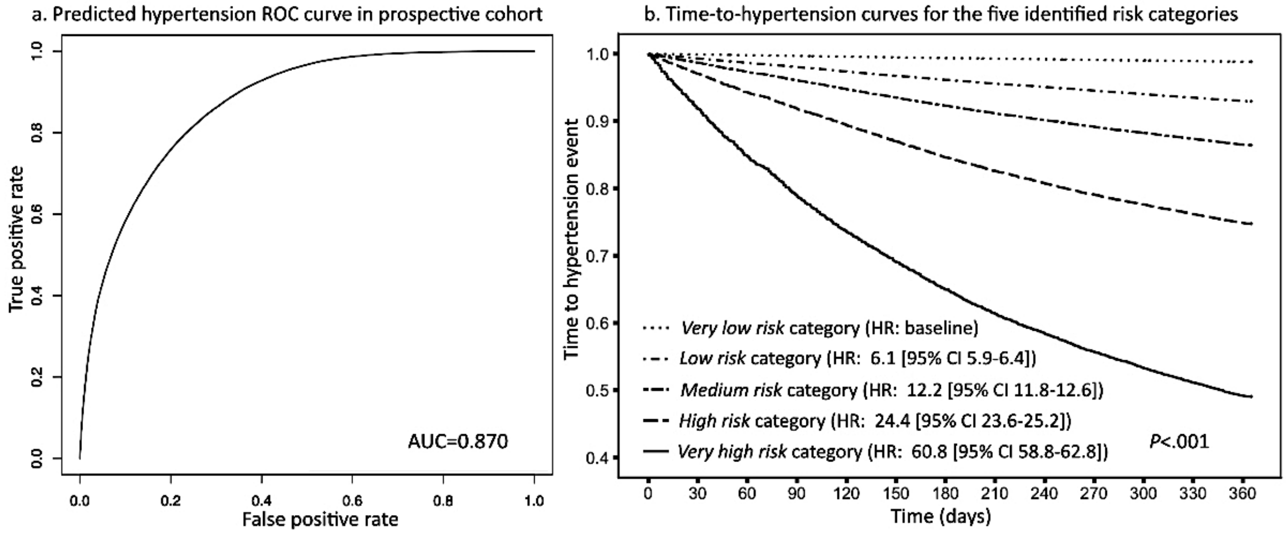


Figure 3. Six classifications of the 169 electronic health record (EHR)–based impactful features recognized by our risk model. COPD: chronic obstructive pulmonary disease; CKD: chronic kidney disease.

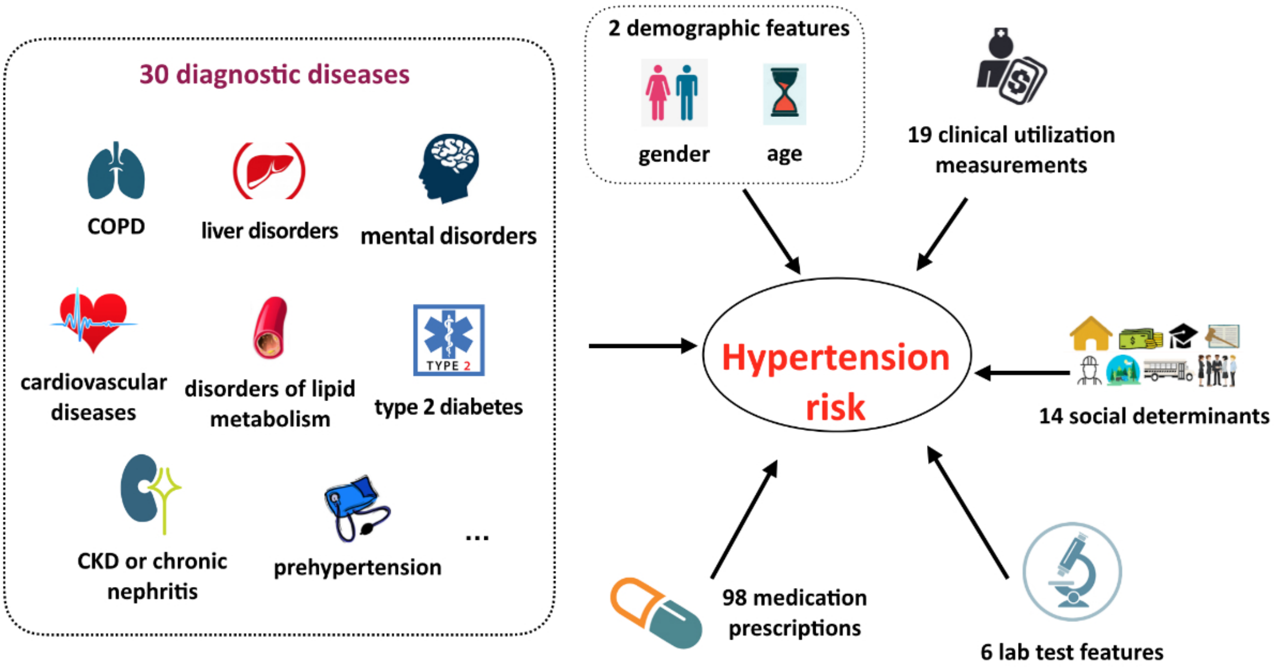
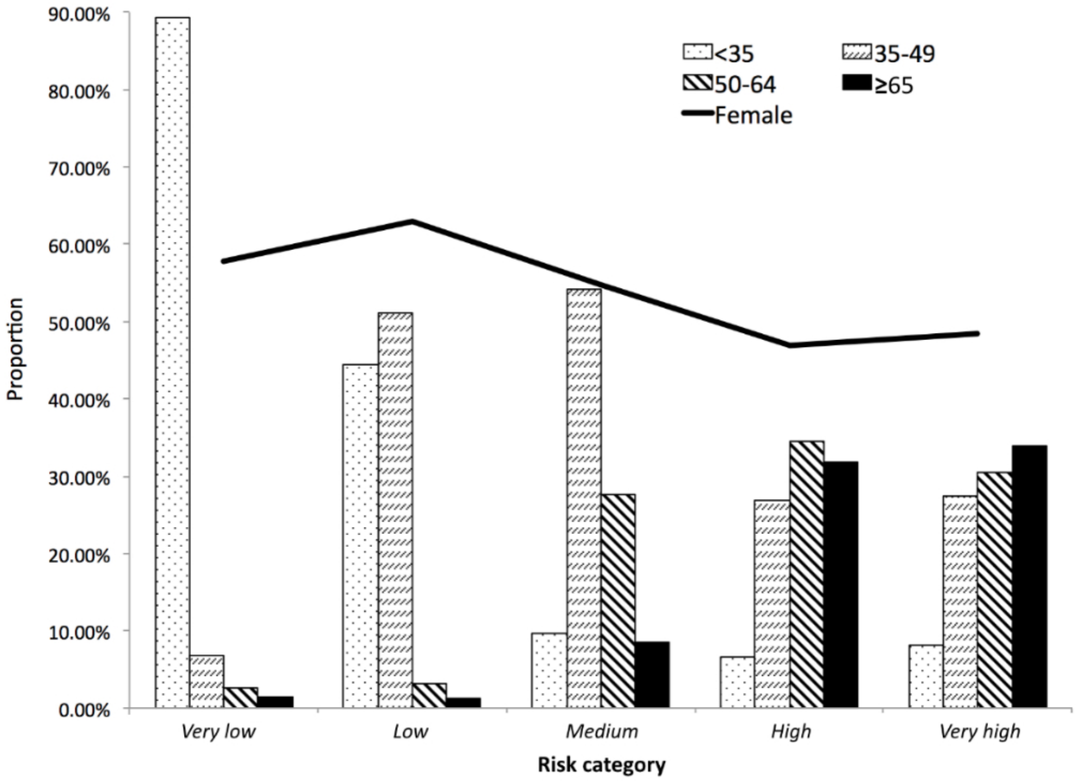


Figure 4. Constituent ratios of age and gender (female) subgroups across the identified five risk categories. Age groups (years): <35, 35-49, 50-64, and ≥65.



To maintain the consistency of this feature, other ages were aggregated into categories as well. In addition, this age categorization could help us capture natural differentials in age-related socioeconomic backgrounds such as work or retirement status and eligibility for pensions and medicare and investigate their impact on hypertension risk and incidence [36]. Therefore, from the *very low risk* to the *very high risk* category of hypertension, constituent ratios were calculated for four distinct age groups and compared with each other (Figure 4). Low-risk categories consisted of young people, whereas high-risk categories were concentrated with older individuals, confirming age group as a strong predictor of hypertension. Younger individuals (<35 years) comprised almost 89.27% (340,621/381,544) of the *very low risk* category. Middle-aged adults (35-49 years) were overwhelmingly dominant in the *low risk* and *medium risk* categories, with proportions of 51.04% (53,372/104,565) and 54.11% (53,792/99,415), respectively. On the contrary, individuals in the age range of 50 to 64 years occupied 34.57% (18,651/53,957) of the *high risk* population, and people aged ≥65 years constituted the largest subset (33.88%, 14,000/41,329) of the *very high risk* population (see Multimedia Appendix 4). Following that, we formed a subpopulation with people ≥65 years and found that our hypertension prediction model reached an acceptable to good discriminative ability with this subpopulation, with an AUC of 0.744 (see Multimedia Appendix 5). On the other hand, females were likely to have lower risk of hypertension than males, comprising majorities (58.25%, 341,072/585,524) of the three relatively low-risk categories (ie, *very low*, *low*, and *medium*) but decreasing to proportions of 46.89% (25,300/53,957) and 48.41% (20,008/41,329) in the *high risk* and *very high risk* groups (Figure 4). In summary, the *very high risk* population mainly comprised aged (>50 years) males, who occupied almost one-third of this category.

Diagnosed Chronic Diseases

Most individuals in the *very high risk* population had a history of multiple chronic medical conditions. Most commonly CVDs, type 2 diabetes, and disorders of lipid metabolism affected 25.08% (10,367/41,329), 20.17% (8,336/41,329), and 8.76% (3619/41,329) of the population, respectively (see Multimedia Appendix 4), compared with only 0.65% (2473/381,544), 0.26% (1006/381,544), and 0.27% (1036/381,544) of the *very low risk* cohort.

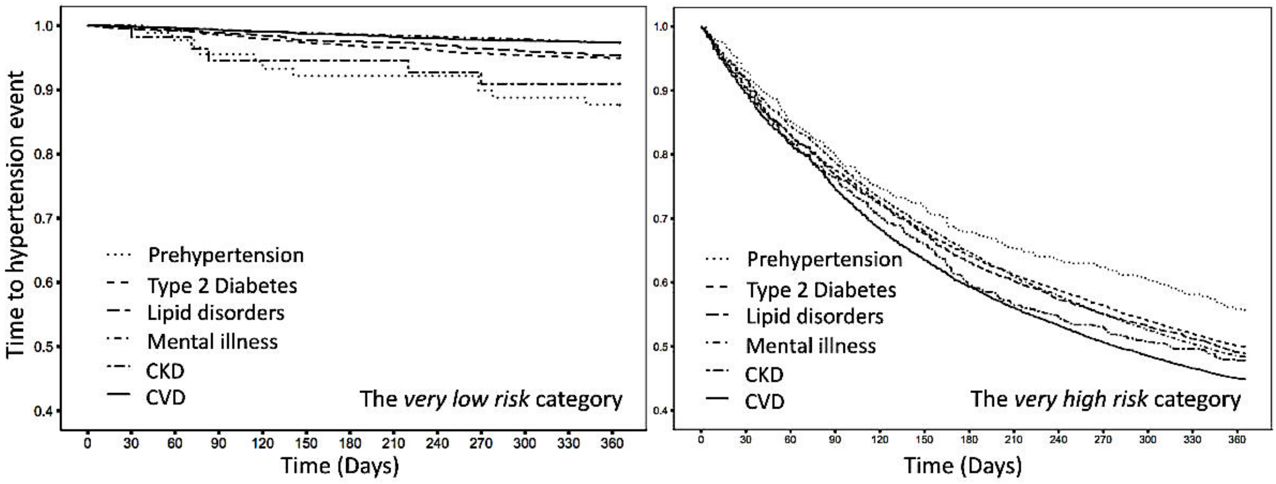
To further explore real-time hypertension events in the *very high risk* category, we investigated time-to-hypertension curves in terms of disease subgroups (ie, prehypertension, type 2 diabetes, lipid disorders, mental illness, CKD, and CVDs) and compared them with curves generated from the *very low risk* cohort (Figure 5). In the *very high risk* category, 55.09% (5711/10,367) patients with CVD would receive diagnoses of hypertension in the next 1 year, and this probability stayed around 50% for patients with conditions of either type 2 diabetes, lipid disorders, mental illness, or CKD. Thus, a half of individuals with a chronic disease in the *very high risk* category would develop hypertension within the next 1 year, with CVD patients at the highest risk of hypertension during the next 1 year, whereas prehypertension patients had the lowest risk, implying different causal mechanism of hypertension under different disease scenarios at the *very high risk* stage. In contrast, in the *very low risk* category, more than 85% of patients were free from development of hypertension in the next 1 year, regardless of chronic conditions, where the survival curve dropped fastest for patients with prehypertension but went down slowly for the cardiovascular subgroup, opposite to the trends at the *very high risk* stage.

Mental Illness

Apart from medications for CVD, type 2 diabetes, lipid disorders, and COPD, our prediction model also recognized 18 important medications prescribed for mental health diagnoses as powerful predictors of 1-year hypertension risk, which were mainly drugs prescribed for mood disorders such as depression, anxiety disorders, and schizophrenia disorders (see Multimedia Appendix 6). In our original HIE dataset from Maine, patient records related to mental health were initially blocked because of privacy protection issues. Therefore, our study utilized these consumed drugs as proxies to inspect the association between mental illness and risk of hypertension. People under treatment for depression, anxiety, and schizophrenia disorders were significantly enriched in the *very high risk* group for hypertension, with proportions of 22.21% (9181/41,329), 13.40% (5538/41,329), and 2.71% (1119/41,329), respectively, but much less prevalent in the *very low risk* group (see Multimedia Appendix 7).

Taking that most prevalent depression as an example, we further divided the prospective cohort into two subgroups of individuals who did and did not carry other chronic disease diagnoses and validated the association between mental disorders and hypertension risk accordingly.

Figure 5. Kaplan-Meier curves depicting time-to-hypertension for the next 1 year according to coexisting disease subgroup for the very low risk (left) and very high risk (right) categories of the prospective cohort. Disease subgroups comprised patients who received diagnoses of either prehypertension, type 2 diabetes, lipid disorders, mental illness, chronic kidney disease (CKD), or cardiovascular disease (CVD), respectively.



By using multivariable Cox proportional hazards models, the contribution of depression to the hazard for hypertension was assessed after adjustment for age and gender (Figure 6). In the subgroup that comprised people with other chronic conditions, depression status revealed a relatively small effect on the 1-year hypertension hazard (HR: 1.1 [95% CI 1.1-1.2]), lower than both age and gender. However, in the subgroup of patients without other chronic physical conditions, we found that depression was associated with a two-fold hypertension risk (HR: 2.0 [95% CI 1.9-2.0]). Following that, time-to-hypertension curves were stratified by age (<65 and ≥65 years), depression status, and gender, in order of priority. Age, as the most impactful predictor, attained the highest HR (HR: 2.0 [95% CI 4.6-4.8]), whereas depression status was the second impactful feature. Therefore, as depicted in the two lowest survival curves in Figure 6, for individuals ≥65 years (mostly assembled in the *very high risk* category), a diagnosis of depression would double the 1-year hazard of hypertension even if there were no other chronic conditions, regardless of gender.

Clinical Utilization Indicators

More severe health conditions are expected to consume more resources. Accordingly, clinical utilization indicators in our study revealed similar patterns to those highlighted diseases when compared across five risk categories. These utilization parameters such as outpatient visits, inpatient admissions, clinical cost, total number of consumed prescriptions, and total number of laboratory tests in the past year were lowest in the *very low risk* group but gradually increased from low to high risk groups, with the highest values in the *very high risk* population (see Multimedia Appendix 4).

On the basis of these findings, we focused compared patients in the *very high risk* and *very low risk* categories with respect to average clinical costs in the prior 12 months for 16 chronic disease subgroups, coordinated by their averaged number of chronic diseases (Figure 7). In the *very low risk* category, patients with prehypertension, CVDs, type 2 diabetes, lipid disorders, and mental disorders had relatively low prevalence (ie, small balls in Figure 7) and clustered in the lower left corner (ie, the low-cost and low-chronic-complexity area), very close to the largest bubble representing the *reference* subgroup formed by nonchronic disease individuals. However, in the *very high* hypertension-risk category, people usually suffered from multiple chronic conditions, resulting in more disease comorbidities and more clinical costs, bubbling toward the upper right corner. Moreover, in the *very high risk* category, the clinically costly CKD and chronic nephritis were recognized as the most severe chronic diseases on average, accompanied by six other chronic conditions, whereas CVD, type 2 diabetes, lipid disorders, and mental disease stayed in a moderate level of disease comorbidity and clinical cost.

Social Economic Features

Several interesting results were also found in socioeconomic features. In this study, most socioeconomic features were derived from ZIP code or county-based census and USDA data, and thus, they were recognized as community-level social and environmental indicators. The association between significant socioeconomic factors and the summarized hypertension risk scores were further investigated by Spearman rank correlations (Figure 8). Educational disparities revealed a unique and significant correlation with hypertension risk scores in our prediction model. The percentage of the low-education population (ie, a combination of the 18-24 year old population with less than high school graduate diploma and ≥25-year-old population with less than 12th grade diploma in the area) gradually increased from low- to high-risk categories and ultimately displayed a positive correlation (ρ=.05) with hypertension risk, occupying 8.73% (33,310/381,544) and 9.20% (3804/41,329) of the *very low risk* and the *very high risk* subsets (see Multimedia Appendix 8). In contrast, the averaged proportion of high-educated people that received college or associate’s degree or bachelor’s and higher degree grew in an opposite direction and showed a negative correlation (ρ=−.072) with hypertension risk, by attaining its highest percentage of 49.05% (187,148/381,544) in the *very low risk* group and decreased to its lowest value of 47.85% (19,776/41,329) in the *very high risk* population (see Multimedia Appendix 8).

Figure 6. Predicted time-to-hypertension curves for patients (a) who received no diagnosis of other chronic diseases except possible depression and (b) who had at least one diagnosis of other chronic disease. Curves for both subgroups were stratified by age (<65 vs ≥65 years), gender (male vs female), and depression diagnosis (depression vs control).

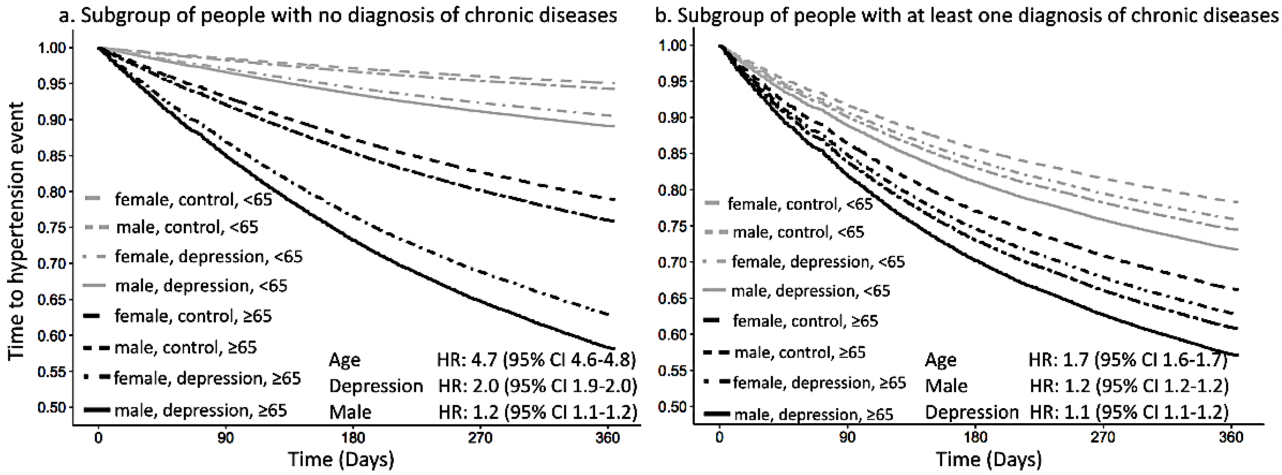


Figure 7. Patients' average clinical costs in the past 12 months against the average number of chronic diseases. The balls were formed by 16 common disease subgroups under the very low risk (blue balls) and very high risk (yellow balls) categories, respectively. The ball size indicates the proportion of the disease subgroup under this risk category. The 16 chronic diseases were prehypertension, cardiovascular disease (CVD), type 2 diabetes, lipid disorders, chronic kidney disease (CKD), chronic nephritis, depression, anxiety, schizophrenia, chronic obstructive pulmonary disease (COPD), liver disorders, cancer, anemia, hypokalemia, idiopathic hypersomnia, and hyposmolality or hyponatremia. Reference groups consisted of patients with no diagnosis of any above chronic diseases.

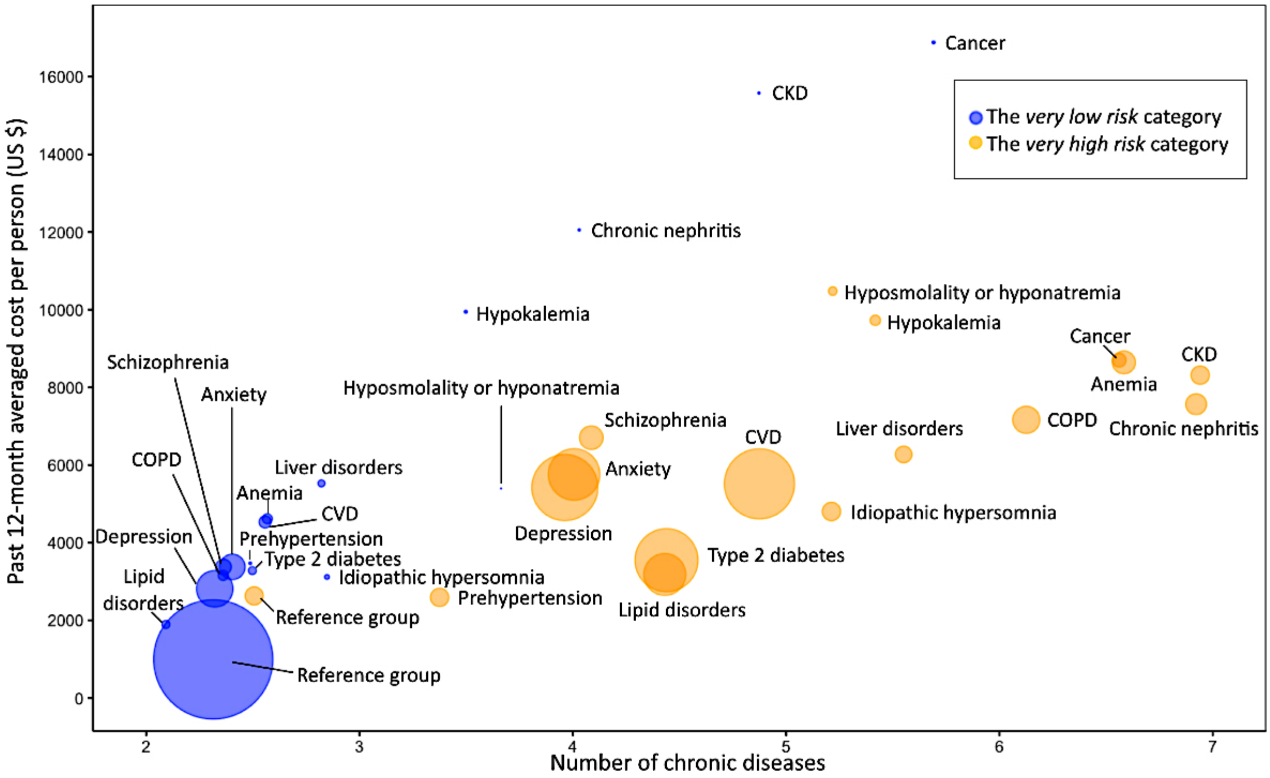
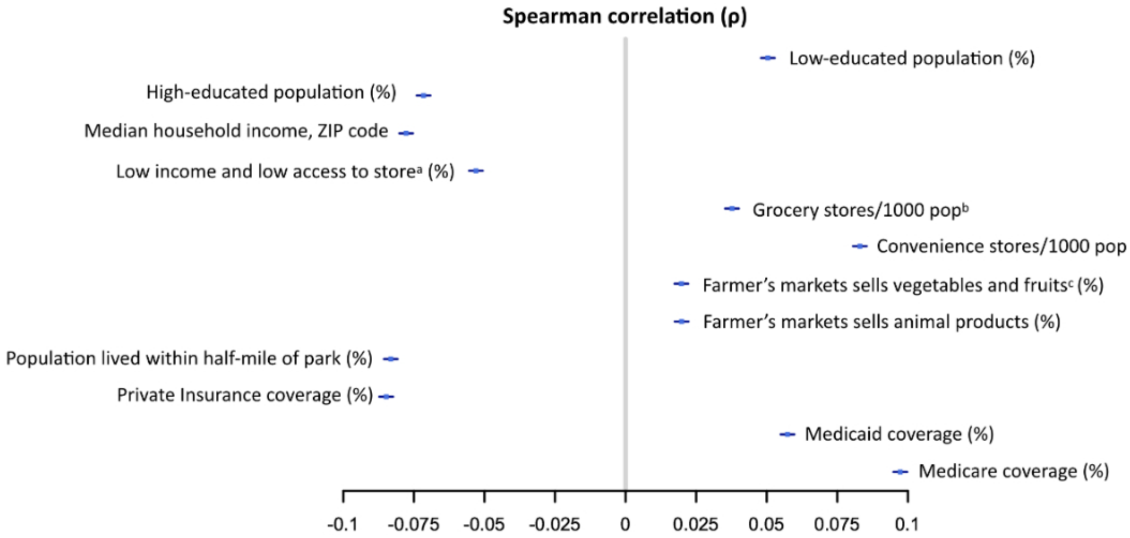


Figure 8. Spearman rank correlation between 12 socioeconomic features and prospective hypertension risk scores. (a) Percentage of people in a county with low income and living more than 1 mile from a supermarket or large grocery store if an urban area, or more than 10 miles from a supermarket or large grocery store if a rural area. (b) The number of supermarkets and grocery stores in the county per 1000 county residents. (c) Percentage of farmers' markets in the county that sell fresh fruits and vegetables.



More disparities were found in other social and economic indicators, both community-level persecutors of and defenders against hypertension (Figure 8). Those risk indicators with positive correlations (ρ>0) mainly comprised measures of nearby available food stores (eg, number of grocery and convenience stores per 1000 population and percentages of farmers' markets selling fruit and vegetables and animal products within a certain region) and the percentage of people benefiting from Medicaid and Medicare type health insurance. On the other hand, the population-based defenders of hypertension (ρ<0) were identified as the median household income, percentage of people with low income and living far from supermarket or large grocery store (>1 mile for an urban area and >10 mile for a rural area), percentage of the population living within a half-mile of

NAVTEQ-based parks, and percentage of the population covered by private health insurance, as the higher those values were, the lower the risk of hypertension would be in the population. In summary, in terms of social determinants, the *very high risk* population was likely to be occupied by low-income and low-educated people who benefited from public insurance and lived in an area near food stores but far from parks.

Discussion

Summary of Main Findings

In this study, we have prospectively validated a risk prediction model of future 1-year incident essential hypertension using EHR data derived from more than 1.5 million people in the state of Maine. The model achieved 0.917 and 0.870 predictive accuracy in retrospective and prospective (validation) cohorts, respectively. On the basis of the risk model, patients were ultimately stratified into five distinct risk categories, ranking the hypertension risk as *very low*, *low*, *medium*, *high*, or *very high*. In the prospective (validation) cohort, these five identified risk categories showed distinct HRs for incident hypertension within the next 1 year (Figure 2), indicating our model's ability to target those most at risk for subsequent prevention management.

Unlike traditional prediction methods, our study adopted a machine learning algorithm called XGBoost for feature selection and model construction. This supervised machine learning technique is designed to discover statistical patterns in high-dimensional and multivariate datasets and is able to handle nonlinear correlations and random errors both in input features and the output variable [28].

Variable correlation is another common issue when using a large number of variables spontaneously to construct a model. For instance, in our study, potential spatial correlation may exist among demographic or socioeconomic features. Most regression algorithms such as generalized linear model assume feature independence and could become less accurate if variable correlation or multicollinearity exist. On the contrary, decision trees (including XGBoost) are nonparametric algorithms [37] that do not assume a functional relationship between outcome and features as is required by linear regression models. They naturally perform a greedy algorithm of finding the best splits in the data that maximize the entropy reduction of the outcome during each split. Therefore, once a feature is chosen by XGBoost, the importance of any highly correlated feature will be significantly reduced as the effective split is already achieved by the original feature; accordingly, correlated features will no longer effectively reduce the entropy of the outcome. As a result, XGBoost and other decision tree algorithms are robust to correlated features.

As a result, XGBoost provided a more accurate prediction model (AUC of 0.87 in the prospective cohort) than prior models with no extra cost by capturing previously ignored but potentially powerful predictors from patients' current health conditions, chronic disease and medication history, clinical utilization measures, and social determinants. Specifically, individuals with the highest risk in our study were typically older (>50 years) and suffered from multiple chronic conditions (concurrence of type 2 diabetes, lipid disorders, mental illness, cardiovascular disease, etc). In the socioeconomic domain, low-income and low-educated people who had public insurance and lived near food stores but far away from parks were enriched in the *very high risk* category. In the prospective cohort, 50.93% (21,050/41,329) of people in the *very high risk* group were diagnosed with hypertension during the next 1 year. Such a high incidence of hypertension in the *very high risk* subgroup made us believe that a short-term (1 year) prediction of hypertension was necessary to identify a high-risk cohort, as well as promote their follow-up intervention or prevention.

Interpretation of Meaningful Risk Predictors and its Implications for Prevention and Early Intervention

Multiple Chronic Conditions

Most individuals in the *very high risk* group suffered from other medical conditions, and only 1.56% (645/41,329) had no *other* chronic medical conditions, whereas the majority (73.10%, 278,923/381,544) of the *very low risk* population had no *other* chronic disease diagnosis. Moreover, as the number of people with diagnostic chronic diseases increased, those with multiple chronic conditions (MCCs), also known as concurrent chronic conditions [38], became very common in the *very high* risk category (Figure 7), occupying almost a third (29.60%, 12,234/41,329) of the population, which also led to a dramatically increased burden of health care utilization and cost (Figure 5). In recent years, epidemiological studies indicated the high-level and continually increased prevalence of MCC worldwide. About a third of the US population suffer from MCC, the majority of which are working-age adults (45-64 years), whereas the prevalence is almost 80% among people ≥65 years [39,40]. Consistent with these findings, our *very high* hypertension risk category could be verified as an MCC population mainly consisting of elderly people with comorbidity of multiple chronic diseases. Hypertension is usually recognized as a major risk factor for cardiovascular and renal diseases. Conversely, within a complex reciprocal interrelationship, both essential and secondary hypertension can also be a consequence of chronic conditions [41,42], especially for elderly people [43]. Therefore, based on our prediction model, high risk of hypertension can be seen as the consequence of a long-term interplay of multiple chronic conditions via possible mechanisms such as polyunsaturated omega-3 free fatty acid deficiency, vascular endothelial dysfunction, immune dysregulation, and unhealthy lifestyle [44,45].

In addition, the definition of prehypertension varied in different studies from time to time [46-48]. In the sixth report of the Joint National Committee (JNC), prehypertension was defined as a systolic pressure of 130 to 139 mm Hg or a diastolic pressure of 85 to 89 mm Hg, whereas the updated seventh report of the JNC (JNC 7) defines prehypertension as a systolic pressure of 120 to 139 mm Hg or a diastolic pressure of 80 to 89 mm Hg. In our study, we derived the feature prehypertension from the ICD-9-CM category of 796.2 and the ICD-10-CM category of R03.0, defined according to JNC 7. Therefore, caution is needed when comparing our findings with those of previous studies, considering the different definitions of prehypertension.

Mental Illness

The strength of association between mental disorders and subsequent onset of hypertension has been of interest for a long time. Much work has been done on specific psychological domains of depression, anxiety, impulsive eating disorders, and other mental disorders [49,50]. Furthermore, several studies have focused on the effects of mental disorders on subsequent hypertension [51,52]. Recognizing the impact of mental disorders on the onset of hypertension is clinically meaningful for several reasons. First, both the strength of effect and potential mechanisms of psychological and mental disorders to cause physical health conditions remain unclear [53,54]. Second, such association studies are driven by the ongoing worldwide increase of morbidity related to mental disorders.

In our study, the 1-year hypertension risk model and the depression-based survival analysis provided strong evidence of such an association, especially for individuals having no other chronic physical conditions. That is, about 32.17% (1226/3811) of elderly (≥65 years) patients having no chronic conditions but undergoing treatment for depression would develop hypertension within the next 1 year. EHR data in the state of Maine coding mental health diagnoses were masked for privacy protection; only consumed mental disorder drugs could be accessed as proxies. Therefore, as a limitation of the study, we cannot directly ascertain the effect of mental health disorders or the treatment thereof on the progression of hypertension. However, we can examine the relationship between mental illness diagnoses or therapeutics and hypertension development in several different ways according to the large amount of previous studies exploring this relationship [49-52,55-59]. First, mental illness itself could be a risk factor of hypertension, where possible mechanisms could relate to internal responses from altered sleep patterns, sympathoadrenal hyperreactivity, various neurotransmitter abnormalities, or altered inflammatory processes [55,56], as well as exposure to similar risk factors (eg, early childhood adversity) [57]. It is reported that depression can trigger deregulation in the sympathetic nervous system and hypothalamic-pituitary-adrenal axis, causing severe consequences such as metabolic syndrome and elevated risk of hypertension [59]. Second, mental illness medication could also independently increase the risk of hypertension through various drug-related side effects, which are connected to increasing rates of unhealthy lifestyles such as smoking, alcohol intake, physical inactivity, and drug addiction [45]. Certain drug treatments could also lead to the reprogramming of hypertension-related metabolic pathways. Previous cohort studies have found an increase in hypertension risk when tricyclic antidepressants are used; a possible mechanism is the effect of the aforementioned antidepressants on vagal control over the heart [58]. Future work is needed to address the causal relationship between mental illness and hypertension.

Social Determinants

Social determinants of health are defined as “the structural determinants and conditions in which people are born, grow, live, work, and age” [60,61] and include social or environmental supportive factors of health, such as socioeconomic status, education level, employment and income, the physical environment, and access to health care. It has been reported that health disparities and inequalities in these social factors, especially education and poverty, accounted for over a third of total deaths in the United States [62]. Social determinants of health could be collected from global, national, and community levels, manifesting varying localized policies and actions on health promotion. Consistent with previous studies [63-66], our model demonstrated that the hazard of 1-year incident hypertension increased as the median household income and the percentage of high-educated population within a ZIP code area went down. In fact, lower education level could be translated as one of the initial driving and predisposing social features, leading to disparities of subsequent incident hypertension. That is, low education level could directly affect family income and ultimately increase the likelihood of having public health insurance, such as Medicaid or Medicare. In addition to those common social determinants of health, some lifestyle-related social factors were also identified in our study, namely, indicators of nearby accessible food stores and parks. The location of food stores and the types of products sold could affect underlying unhealthy diet habits, whereas environmental supportive facilities such as the concentration of parks or playgrounds in the living area may directly shape people’s physical activities and ultimately decrease hypertension risk. Another interesting finding was that, although low income level would increase hypertension risk overall, lower income populations had reduced hypertension risk if they lived far from food or grocery stores, implying an independent and critical impact of food consumption habits on hypertension regardless of income level. As the majority of the Maine population is white, our study failed to find any significant racial and ethnic disparities for hypertension risk, and it is possible that that our model is not as accurate with regard to individuals of other race or ethnicity. Another limitation of our study is that such community-level social determinants may potentially be less precise and could leave some useful social determinants uncaptured in the risk model.

Application of the Risk Model

Beyond the prediction of hypertension risk itself, subsequent actions to prevent and intervene are necessary in high-risk patients. In our *very high risk* category, 21.26% (8788/41,329) of the population would receive a diagnosis of hypertension within the first 3 months of next year, and this number would gradually increase to 50.93% (21,050/41,329) after the 1-year period. Therefore, monitoring those high-risk patients as soon as possible and developing a personalized longitudinal intervention plan is important to prevent or delay the development of incident hypertension, as well as to reduce corresponding health care expenditures. Given the fact that 98.44% (40,684/41,329) people in the *very high risk* cohort had other chronic diseases, one possible early intervention plan could be more active monitoring and treatment of relevant diseases, blocking possible pathways from chronic diseases to hypertension development. For instance, strong evidence has illustrated that patients with depressive disorders can be effectively treated in primary care by using either low-cost remote cognitive behavioral therapy or face-to-face psychological interventions, which may ultimately benefit the prevention of hypertension [67]. Moreover, it is also illustrated in our study that a high concentration of parks or playgrounds in the living area can reduce hypertension risk, most likely by shaping people physical activities, implying the importance of community level intervention such as increased supportive neighborhood environments to advocate more healthy lifestyles.

Study Limitations

Our study has several limitations. First, missing data is inevitable in EHR records, and thus, data organization and extraction are critical. The KNN approach was used in our data preprocessing for missing data imputation. One limitation of this approach is that, when most variables are missing for a certain patient, the KNN cannot be identified appropriately, and therefore, the imputation can be less accurate and cause bias [68]. Second, by directly using EHR data, it is possible that some patients with hypertension failed to have a record of this diagnosis, leading to an underestimation of hypertension prevalence in the study. Third, as the primary and secondary diagnoses and procedures were coded in ICD-10-CM after October 1, 2015, we mapped them back to ICD-9-CM using the GEM tool to keep the data consistent. This mapping strategy may be inadequate as ICD-10 has more than 65,000 codes, whereas ICD-9-CM has only 13,000. Immediately after the ICD-10-CM transition on October 1, 2015, the large number of codes and the increased specificity could have been potential challenges for physicians in terms of code assignment, leading to potential systematic bias [69]. Fourth, some traditional risk factors could not be directly captured by structured EHR data. Future natural language processing [70] could utilize the unstructured EHR data to extract additional risk factors including individual-level lifestyle information (eg, diet habit and physical activity).

Conclusions

In summary, we have constructed and prospectively validated a risk prediction model of future 1-year incident essential hypertension using EHR data derived from more than 1.5 million people in the state of Maine. The EHR-based model achieved 0.917 and 0.870 predictive accuracy in retrospective and prospective (validation) cohorts, respectively, and successfully stratified patients into five distinct risk categories from *very low*, *low*, *medium*, *high*, to *very high*. Our real-time predictive analytic model has already been deployed in the state of Maine. Integration of such predictive analysis into clinical prescriptive solutions may help health care providers target high-risk populations, tailor the prescription and intensity of treatment solutions to such high-risk cohorts, improve decision making and patient adherence to prescribed intervention, and eventually benefit individuals' health and quality of life while reducing health care costs.

Acknowledgments

The authors would like to thank and express their gratitude to the hospitals, medical practices, physicians, and nurses participating in Maine's HIE. They also thank the biostatistics colleagues at the Department of Health Research and Policy for critical discussions. CY is supported by the National Natural Science Foundation of China (Award #81402762).

Authors' Contributions

CY, TF, YZ, SH, BJ, MX, ML, XZ, QW, OW, YG, CZ, and YL carried out the initial analysis and interpretation of data and drafted the initial manuscript. FS, KGS, EW, DBM, and XBL conceptualized and designed the study and critically reviewed and revised the manuscript. DSC and STA coordinated and supervised data acquisition and critically reviewed and revised the manuscript. All authors have read and approved this submission for publication. All authors have agreed to be accountable for all aspects of the work.

Conflicts of Interest

KGS, EW, and XBL are cofounders and equity holders of HBI Solutions, Inc, which is currently developing predictive analytics solutions for health care organizations. The research and research results are not, in any way, associated with Stanford University. There are no patents, further products in development, or marketed products to declare. This does not alter our adherence to all the journal policies on sharing data and materials, as detailed online in the guide for authors.

Multimedia Appendix 1

List of social determinant variables downloaded from the US census and United States Department of Agriculture (USDA) websites, detailed in the data source and mapping method.

[PDF File (Adobe PDF File), 41KB - jmir_v20i1e22_app1.pdf]

Multimedia Appendix 2

The performance of the 1-year hypertension risk prediction model in the prospective cohort, summarized in PPV, sensitivity, and specificity.

[PDF File (Adobe PDF File), 19KB - jmir_v20i1e22_app2.pdf]

Multimedia Appendix 3

The most impactful 80 features selected by our hypertension prediction model.

[PDF File (Adobe PDF File), 36KB - jmir_v20i1e22_app3.pdf]

Multimedia Appendix 4

Distribution of impactful risk predictors across the five risk categories, specified as features of demographics, diagnosed diseases, and clinical utilization.

[PDF File (Adobe PDF File), 25KB - jmir_v20i1e22_app4.pdf]

Multimedia Appendix 5

ROC curves and AUC values of subgroups.

[PDF File (Adobe PDF File), 134KB - jmir_v20i1e22_app5.pdf]

Multimedia Appendix 6

Distribution of impactful medications of depression, anxiety, and schizophrenia across the five risk categories.

[PDF File (Adobe PDF File), 24KB - jmir_v20i1e22_app6.pdf]

Multimedia Appendix 7

Constituent ratios of three mental diseases (depression, anxiety, and schizophrenia) in the very low risk and very high risk categories.

[PDF File (Adobe PDF File), 136KB - jmir_v20i1e22_app7.pdf]

Multimedia Appendix 8

Distribution of social determinant indicators across the five categories with very low, low, medium, high, and very high risk of hypertension.

[PDF File (Adobe PDF File), 28KB - jmir_v20i1e22_app8.pdf]

References

1. Benjamin EJ, Blaha MJ, Chiuve SE, Cushman M, Das SR, Deo R, American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart Disease and Stroke Statistics-2017 Update: A Report from the American Heart Association. Circulation 2017 Mar 7;135(10):e146-e603. [doi: 10.1161/CIR.0000000000000485] [Medline: 28122885]
2. World Health Organization. Blood Pressure URL: http://www.who.int/gho/ncd/risk_factors/blood_pressure_prevalence/en/ [accessed 2017-10-23] [WebCite Cache ID 6wAurqifr]
3. Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo Jr JL, National Heart, Lung, and Blood Institute Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure, National High Blood Pressure Education Program Coordinating Committee. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. J Am Med Assoc 2003 May 21;289(19):2560-2572. [doi: 10.1001/jama.289.19.2560] [Medline: 12748199]
4. Elliott WJ. The economic impact of hypertension. J Clin Hypertens (Greenwich) 2003;5(3 Suppl 2):3-13 [FREE Full text] [Medline: 12826765]
5. Vasan RS, Larson MG, Leip EP, Kannel WB, Levy D. Assessment of frequency of progression to hypertension in non-hypertensive participants in the Framingham Heart Study: a cohort study. Lancet 2001 Nov 17;358(9294):1682-1686. [doi: 10.1016/S0140-6736(01)06710-1] [Medline: 11728544]
6. Effects of weight loss and sodium reduction intervention on blood pressure and hypertension incidence in overweight people with high-normal blood pressure. The Trials of Hypertension Prevention, phase II. The Trials of Hypertension Prevention Collaborative Research Group. Arch Intern Med 1997 Mar 24;157(6):657-667. [Medline: 9080920]
7. He J, Whelton PK, Appel LJ, Charleston J, Klag MJ. Long-term effects of weight loss and dietary sodium reduction on incidence of hypertension. Hypertension 2000 Feb;35(2):544-549 [FREE Full text] [Medline: 10679495]
8. Julius S, Nesbitt SD, Egan BM, Weber MA, Michelson EL, Kaciroti N, Trial of Preventing Hypertension (TROPHY) Study Investigators. Feasibility of treating prehypertension with an angiotensin-receptor blocker. N Engl J Med 2006 Apr 20;354(16):1685-1697. [doi: 10.1056/NEJMoa060838] [Medline: 16537662]

9. Parikh NI, Pencina MJ, Wang TJ, Benjamin EJ, Lanier KJ, Levy D, et al. A risk score for predicting near-term incidence of hypertension: the Framingham Heart Study. Ann Intern Med 2008 Jan 15;148(2):102-110. [Medline: 18195335]
10. Echouffo-Tcheugui JB, Batty GD, Kivimäki M, Kengne AP. Risk models to predict hypertension: a systematic review. PLoS One 2013 Jul 5;8(7):e67370 [FREE Full text] [doi: 10.1371/journal.pone.0067370] [Medline: 23861760]
11. Singh A, Nadkarni G, Gottesman O, Ellis SB, Bottinger EP, Guttag JV. Incorporating temporal EHR data in predictive models for risk stratification of renal function deterioration. J Biomed Inform 2015 Feb;53:220-228 [FREE Full text] [doi: 10.1016/j.jbi.2014.11.005] [Medline: 25460205]
12. National Institutes of Health. 2017 Mar 24. Is the EHR the New Big Data? URL: https://datascience.nih.gov/BlogIsTheEHR [accessed 2017-11-19] [WebCite Cache ID 6v9ULMIuP]
13. Hersh WR. Adding value to the electronic health record through secondary use of data for quality assurance, research, and surveillance. Am J Manag Care 2007 Jun;13(6 Part 1):277-278 [FREE Full text] [Medline: 17567224]
14. Goldstein BA, Navar AM, Pencina MJ, Ioannidis JP. Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review. J Am Med Inform Assoc 2017 Jan;24(1):198-208. [doi: 10.1093/jamia/ocw042] [Medline: 27189013]
15. Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLoS One 2017 Apr 4;12(4):e0174944 [FREE Full text] [doi: 10.1371/journal.pone.0174944] [Medline: 28376093]
16. Li L, Cheng WY, Glicksberg BS, Gottesman O, Tamler R, Chen R, et al. Identification of type 2 diabetes subgroups through topological analysis of patient similarity. Sci Transl Med 2015 Oct 28;7(311):311ra174 [FREE Full text] [doi: 10.1126/scitranslmed.aaa9364] [Medline: 26511511]
17. Gultepe E, Green JP, Nguyen H, Adams J, Albertson T, Tagkopoulos I. From vital signs to clinical outcomes for patients with sepsis: a machine learning basis for a clinical decision support system. J Am Med Inform Assoc 2014;21(2):315-325 [FREE Full text] [doi: 10.1136/amiajnl-2013-001815] [Medline: 23959843]
18. Rana S, Tran T, Luo W, Phung D, Kennedy RL, Venkatesh S. Predicting unplanned readmission after myocardial infarction from routinely collected administrative hospital data. Aust Health Rev 2014 Sep;38(4):377-382. [doi: 10.1071/AH14059] [Medline: 25001433]
19. Hao S, Jin B, Shin AY, Zhao Y, Zhu C, Li Z, et al. Risk prediction of emergency department revisit 30 days post discharge: a prospective study. PLoS One 2014 Nov 13;9(11):e112944 [FREE Full text] [doi: 10.1371/journal.pone.0112944] [Medline: 25393305]
20. Himes BE, Dai Y, Kohane IS, Weiss ST, Ramoni MF. Prediction of chronic obstructive pulmonary disease (COPD) in asthma patients using electronic medical records. J Am Med Inform Assoc 2009;16(3):371-379 [FREE Full text] [doi: 10.1197/jamia.M2846] [Medline: 19261943]
21. Hippisley-Cox J, Coupland C. Predicting the risk of chronic kidney disease in men and women in England and Wales: prospective derivation and external validation of the QKidney Scores. BMC Fam Pract 2010 Jun 21;11:49 [FREE Full text] [doi: 10.1186/1471-2296-11-49] [Medline: 20565929]
22. Hippisley-Cox J, Coupland C, Robson J, Brindle P. Derivation, validation, and evaluation of a new QRISK model to estimate lifetime risk of cardiovascular disease: cohort study using QResearch database. BMJ 2010 Dec 9;341:c6624 [FREE Full text] [Medline: 21148212]
23. Hippisley-Cox J, Coupland C. Identifying patients with suspected colorectal cancer in primary care: derivation and validation of an algorithm. Br J Gen Pract 2012 Jan;62(594):e29-e37 [FREE Full text] [doi: 10.3399/bjgp12X616346] [Medline: 22520670]
24. Hippisley-Cox J, Coupland C. Identifying patients with suspected lung cancer in primary care: derivation and validation of an algorithm. Br J Gen Pract 2011 Nov;61(592):e715-e723 [FREE Full text] [doi: 10.3399/bjgp11X606627] [Medline: 22054335]
25. Hippisley-Cox J, Coupland C. Symptoms and risk factors to identify men with suspected cancer in primary care: derivation and validation of an algorithm. Br J Gen Pract 2013 Jan;63(606):e1-10 [FREE Full text] [doi: 10.3399/bjgp13X660724] [Medline: 23336443]
26. Hippisley-Cox J, Coupland C. Identifying patients with suspected gastro-oesophageal cancer in primary care: derivation and validation of an algorithm. Br J Gen Pract 2011 Nov;61(592):e707-e714 [FREE Full text] [doi: 10.3399/bjgp11X606609] [Medline: 22054334]
27. Chen M, Hao Y, Hwang K, Wang L, Wang L. Disease prediction by machine learning over big data from healthcare communities. IEEE Access 2017;5:8869-8879. [doi: 10.1109/ACCESS.2017.2694446]
28. Jensen PB, Jensen LJ, Brunak S. Mining electronic health records: towards better research applications and clinical care. Nat Rev Genet 2012 Jun 2;13(6):395-405. [doi: 10.1038/nrg3208] [Medline: 22549152]
29. Wu J, Roy J, Stewart WF. Prediction modeling using EHR data: challenges, strategies, and a comparison of machine learning approaches. Med Care 2010 Jun;48(6 Suppl):S106-S113. [doi: 10.1097/MLR.0b013e3181de9e17] [Medline: 20473190]
30. ICD.Codes. ICD-10-CM to ICD-9-CM URL: https://icd.codes/convert/icd10-to-icd9-cm [accessed 2017-11-20] [WebCite Cache ID 6vAqCpkFl]

31. The United States Census Bureau. 2017 Oct 25. URL: https://www.census.gov/ [accessed 2017-10-25] [WebCite Cache ID 6uTzRZU0g]
32. United States Department of Agriculture. Food Environment Atlas URL: https://www.ers.usda.gov/data-products/food-environment-atlas.aspx [accessed 2017-10-24] [WebCite Cache ID 6uTzZL4L0]
33. Batista G, Monard MC. A study of k-nearest neighbour as an imputation method. Front Artif Intell Appl 2002;87:251-260.
34. Hao S, Fu T, Wu Q, Jin B, Zhu C, Hu Z, et al. Estimating one-year risk of incident chronic kidney disease: retrospective development and validation study using electronic medical record data from the state of Maine. JMIR Med Inform 2017 Jul 26;5(3):e21 [FREE Full text] [doi: 10.2196/medinform.7954] [Medline: 28747298]
35. Plagnol V, Curtis J, Epstein M, Mok KY, Stebbings E, Grigoriadou S, et al. A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. Bioinformatics 2012 Nov 1;28(21):2747-2754 [FREE Full text] [doi: 10.1093/bioinformatics/bts526] [Medline: 22942019]
36. van Zon SK, Bültmann U, Mendes de Leon CF, Reijneveld SA. Absolute and relative socioeconomic health inequalities across age groups. PLoS One 2015 Dec 30;10(12):e0145947 [FREE Full text] [doi: 10.1371/journal.pone.0145947] [Medline: 26717482]
37. Random Forests. 2017. Understand your dataset with XGBoost URL: https://cran.r-project.org/web/packages/xgboost/vignettes/discoverYourData.html [accessed 2017-11-22] [WebCite Cache ID 6wAwpvOuR]
38. Sambamoorthi U, Tan X, Deb A. Multiple chronic conditions and healthcare costs among adults. Expert Rev Pharmacoecon Outcomes Res 2015;15(5):823-832 [FREE Full text] [doi: 10.1586/14737167.2015.1091730] [Medline: 26400220]
39. US Department of Health and Human Services. 2010. Multiple Chronic Conditions - A Strategic Framework URL: https://www.hhs.gov/sites/default/files/ash/initiatives/mcc/mcc_framework.pdf [accessed 2017-10-25] [WebCite Cache ID 6uU5zRd59]
40. Gerteis J, Izrael D, Deitz D, LeRoy L, Ricciardi R, Miller T, et al. Multiple Chronic Conditions Chartbook URL: https://www.ahrq.gov/sites/default/files/wysiwyg/professionals/prevention-chronic-care/decision/mcc/mccchartbook.pdf [accessed 2017-10-24] [WebCite Cache ID 6uU5lsTCB]
41. Tedla FM, Brar A, Browne R, Brown C. Hypertension in chronic kidney disease: navigating the evidence. Int J Hypertens 2011;2011:132405 [FREE Full text] [doi: 10.4061/2011/132405] [Medline: 21747971]
42. Luke RG. Essential hypertension: a renal disease? A review and update of the evidence. Hypertension 1993 Mar;21(3):380-390 [FREE Full text] [Medline: 8478047]
43. Egan BM, Sutherland SE, Wagner CS, Fleming DO, Davis RA, Bryan ST, et al. Future hypertension guidelines for older americans with multiple chronic conditions. J Am Soc Hypertens 2016 Apr;10(4):e44. [doi: 10.1016/j.jash.2016.03.105]
44. Mayo Clinic. Diseases and Conditions URL: http://www.mayoclinic.org/diseases-conditions/ [accessed 2017-10-24] [WebCite Cache ID 6uU0LedPy]
45. Stapelberg NJ, Neumann DL, Shum DH, McConnell H, Hamilton-Craig I. A topographical map of the causal network of mechanisms underlying the relationship between major depressive disorder and coronary heart disease. Aust N Z J Psychiatry 2011 May;45(5):351-369. [doi: 10.3109/00048674.2011.570427] [Medline: 21500954]
46. The sixth report of the Joint National Committee on prevention, detection, evaluation, and treatment of high blood pressure. Arch Intern Med 1997 Nov 24;157(21):2413-2446. [Medline: 9385294]
47. Vasan RS, Larson MG, Leip EP, Evans JC, O'Donnell CJ, Kannel WB, et al. Impact of high-normal blood pressure on the risk of cardiovascular disease. N Engl J Med 2001 Nov 1;345(18):1291-1297. [doi: 10.1056/NEJMoa003417] [Medline: 11794147]
48. Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL, National Heart, Lung, and Blood Institute Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure, National High Blood Pressure Education Program Coordinating Committee. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. J Am Med Assoc 2003 May 21;289(19):2560-2572. [doi: 10.1001/jama.289.19.2560] [Medline: 12748199]
49. Meng L, Chen D, Yang Y, Zheng Y, Hui R. Depression increases the risk of hypertension incidence: a meta-analysis of prospective cohort studies. J Hypertens 2012 May;30(5):842-851. [doi: 10.1097/HJH.0b013e32835080b7] [Medline: 22343537]
50. Pan Y, Cai W, Cheng Q, Dong W, An T, Yan J. Association between anxiety and hypertension: a systematic review and meta-analysis of epidemiological studies. Neuropsychiatr Dis Treat 2015 Apr 22;11:1121-1130 [FREE Full text] [doi: 10.2147/NDT.S77710] [Medline: 25960656]
51. Grimsrud A, Stein DJ, Seedat S, Williams D, Myer L. The association between hypertension and depression and anxiety disorders: results from a nationally-representative sample of South African adults. PLoS One 2009;4(5):e5552 [FREE Full text] [doi: 10.1371/journal.pone.0005552] [Medline: 19440241]
52. Stein DJ, Aguilar-Gaxiola S, Alonso J, Bruffaerts R, de Jonge P, Liu Z, et al. Associations between mental disorders and subsequent onset of hypertension. Gen Hosp Psychiatry 2014;36(2):142-149 [FREE Full text] [doi: 10.1016/j.genhosppsych.2013.11.002] [Medline: 24342112]
53. Ramasubbu R, Beaulieu S, Taylor VH, Schaffer A, McIntyre RS, Canadian Network for Mood and Anxiety Treatments (CANMAT) Task Force. The CANMAT task force recommendations for the management of patients with mood disorders

and comorbid medical conditions: diagnostic, assessment, and treatment principles. Ann Clin Psychiatry 2012 Feb;24(1):82-90. [Medline: 22303524]

54. Korte SM, Koolhaas JM, Wingfield JC, McEwen BS. The Darwinian concept of stress: benefits of allostasis and costs of allostatic load and the trade-offs in health and disease. Neurosci Biobehav Rev 2005 Feb;29(1):3-38. [doi: 10.1016/j.neubiorev.2004.08.009] [Medline: 15652252]
55. Li Cavoli G, Mulè G, Rotolo U. Renal involvement in psychological eating disorders. Nephron Clin Pract 2011;119(4):c338-41; discussion c341. [Medline: 22135793]
56. Pickering TG. Could hypertension be a consequence of the 24/7 society? The effects of sleep deprivation and shift work. J Clin Hypertens (Greenwich) 2006 Nov;8(11):819-822 [FREE Full text] [doi: 10.1111/j.1524-6175.2006.05126.x] [Medline: 17086023]
57. Stein DJ, Scott K, Haro Abad JM, Aguilar-Gaxiola S, Alonso J, Angermeyer M, et al. Early childhood adversity and later hypertension: data from the World Mental Health Survey. Ann Clin Psychiatry 2010 Feb;22(1):19-28 [FREE Full text] [Medline: 20196979]
58. Licht CM, de Geus EJ, Seldenrijk A, van Hout HP, Zitman FG, van Dyck R, et al. Depression is associated with decreased blood pressure, but antidepressant use increases the risk for hypertension. Hypertension 2009 Apr;53(4):631-638 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.108.126698] [Medline: 19237679]
59. Veith RC, Lewis N, Linares OA, Barnes RF, Raskind MA, Villacres EC, et al. Sympathetic nervous system activity in major depression. Basal and desipramine-induced alterations in plasma norepinephrine kinetics. Arch Gen Psychiatry 1994 May;51(5):411-422. [Medline: 8179465]
60. Marmot M, Friel S, Bell R, Houweling TA, Taylor S, Commission on Social Determinants of Health. Closing the gap in a generation: health equity through action on the social determinants of health. Lancet 2008 Nov 8;372(9650):1661-1669. [doi: 10.1016/S0140-6736(08)61690-6] [Medline: 18994664]
61. Heiman HJ, Artiga S. Kff. 2015. Beyond Health Care: The Role of Social Determinants in Promoting Health and Health Equity URL: https://www.kff.org/disparities-policy/issue-brief/beyond-health-care-the-role-of-social-determinants-in-promoting-health-and-health-equity/ [accessed 2017-10-24] [WebCite Cache ID 6uU5VyonR]
62. Galea S, Tracy M, Hoggatt KJ, Dimaggio C, Karpati A. Estimated deaths attributable to social factors in the United States. Am J Public Health 2011 Aug;101(8):1456-1465. [doi: 10.2105/AJPH.2010.300086] [Medline: 21680937]
63. Morenoff JD, House JS, Hansen BB, Williams DR, Kaplan GA, Hunte HE. Understanding social disparities in hypertension prevalence, awareness, treatment, and control: the role of neighborhood context. Soc Sci Med 2007 Nov;65(9):1853-1866 [FREE Full text] [doi: 10.1016/j.socscimed.2007.05.038] [Medline: 17640788]
64. Baron SL, Steege AL, Marsh SM, Menéndez CC, Myers JR, Centers for Disease Control and Prevention (CDC). Nonfatal work-related injuries and illnesses - United States, 2010. MMWR Suppl 2013 Nov 22;62(3):35-40. [Medline: 24264487]
65. Basu S, Millett C. Social epidemiology of hypertension in middle-income countries: determinants of prevalence, diagnosis, treatment, and control in the WHO SAGE study. Hypertension 2013 Jul;62(1):18-26 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.113.01374] [Medline: 23670299]
66. Kreatsoulas C, Anand SS. The impact of social determinants on cardiovascular disease. Can J Cardiol 2010;26 Suppl C:8C-13C [FREE Full text] [Medline: 20847985]
67. Linde K, Sigterman K, Kriston L, Rücker G, Jamil S, Meissner K, et al. Effectiveness of psychological treatments for depressive disorders in primary care: systematic review and meta-analysis. Ann Fam Med 2015;13(1):56-68 [FREE Full text] [doi: 10.1370/afm.1719] [Medline: 25583894]
68. Rahman SA, Huang Y, Claassen J, Heintzman N, Kleinberg S. Combining Fourier and lagged k-nearest neighbor imputation for biomedical time series data. J Biomed Inform 2015 Dec;58:198-207 [FREE Full text] [doi: 10.1016/j.jbi.2015.10.004] [Medline: 26477633]
69. Rahmathulla G, Deen HG, Dokken JA, Pirris SM, Pichelmann MA, Nottmeier EW, et al. Implementation and impact of ICD-10 (Part II). Surg Neurol Int 2014 Jul 19;5(Suppl 3):S192-S198 [FREE Full text] [doi: 10.4103/2152-7806.137182] [Medline: 25184098]
70. Wang Y, Luo J, Hao S, Xu H, Shin AY, Jin B, et al. NLP based congestive heart failure case finding: a prospective analysis on statewide electronic medical records. Int J Med Inform 2015 Dec;84(12):1039-1047. [doi: 10.1016/j.ijmedinf.2015.06.007] [Medline: 26254876]

Abbreviations

AUC: area under the curve
COPD: chronic obstructive pulmonary disease
CKD: chronic kidney disease
CVD: cardiovascular disease
EHR: electronic health record
GEM: General Equivalence Mapping

HIE:	health information exchange
HR:	hazard ratio
ICD-9-CM:	International Classification of Diseases, 9th Revision, Clinical Modification
JNC:	Joint National Committee
KNN:	k-nearest neighbors
MCC:	multiple chronic conditions
OR:	odds ratio
PPV:	positive predictive value
USDA:	United States Department of Agriculture
ZIP:	zone improvement plan

Edited by G Eysenbach; submitted 26.10.17; peer-reviewed by W Luo, C Fincham, J Wang; comments to author 17.11.17; revised version received 05.12.17; accepted 06.12.17; published 30.01.18.

Please cite as:
Ye C, Fu T, Hao S, Zhang Y, Wang O, Jin B, Xia M, Liu M, Zhou X, Wu Q, Guo Y, Zhu C, Li YM, Culver DS, Alfreds ST, Stearns F, Sylvester KG, Widen E, McElhinney D, Ling X
Prediction of Incident Hypertension Within the Next Year: Prospective Study Using Statewide Electronic Health Records and Machine Learning
J Med Internet Res 2018;20(1):e22
URL: http://www.jmir.org/2018/1/e22/
doi:10.2196/jmir.9268
PMID:29382633

Original Paper

How Experts' Use of Medical Technical Jargon in Different Types of Online Health Forums Affects Perceived Information Credibility: Randomized Experiment With Laypersons

Maria Zimmermann[1], MSc; Regina Jucks[1], Prof Dr

[1] Institute for Psychology in Education, Department of Psychology and Sport Science, University of Münster, Münster, Germany

Corresponding Author:
Maria Zimmermann, MSc
Institute for Psychology in Education
Department of Psychology and Sport Science
University of Münster
Fliednerstraße 21
Münster, 48149
Germany
Phone: 49 251 8339482
Fax: 49 251 8331399
Email: maria.zimmermann@uni-muenster.de

Abstract

Background: Online health forums are widely used, but the quality of advice differs as much as the knowledge backgrounds of the audience members who receive the advice. It is important to understand how people judge the information given online. In line with the communication accommodation theory (CAT), online forums represent specific social contexts of communication which can present either accommodative or nonaccommodative language to an audience. Accordingly, use of accommodative or nonaccommodative language might affect people's perceived trust in the communicator.

Objective: The objective of this study was to investigate how experts who use accommodative (vs nonaccommodative) language are evaluated by passive users of an online forum.

Methods: Participants (n=98) took part in an online experiment and read experts' posts about 10 nutrition myths. Following a 2 x 2 mixed design, experts' posts were written using either low or high amounts of medical technical jargon (MTJ) (within factor) and were directed at different audiences (mainly other medical experts [in a professional forum] vs a user group mainly comprising laypersons [in an advisory forum]) (between factor). Accommodation occurred where experts used high amounts of MTJ to address other medical experts in the professional forum; it also occurred when experts used low amounts of MTJ to address laypersons in the advisory forum. Conversely, nonaccommodation occurred when experts used high amounts of MTJ in the advisory forum and low amounts of MTJ in the professional forum. In each condition, participants evaluated the credibility of the information, the trustworthiness of the experts, and the accommodation by the experts.

Results: Overall, participants judged the credibility of information to be higher when experts used MTJ that was accommodative to the designated audience, $F_{1,95}=3.10$, $P=.04$, $\eta_p^2=.031$. In addition, participants judged the experts in professional forums to be more trustworthy than experts in advisory forums (all $F_{1,96}\geq3.54$, $P \leq.03$, $\eta_p^2\geq.036$). Moreover, participants rated experts who used high amounts of MTJ to have higher competence ($F_{1,96}=37.54$, $P<.001$, $\eta_p^2=.28$], lower integrity ($F_{1,96}=10.77$, $P=.001$, $\eta_p^2=.101$), and lower benevolence ($F_{1,96}=9.75$, $P=.002$, $\eta_p^2=.092$), as well as to have lower perceived accommodation to the audience (all $F_{1,96}\geq72.17$, $P<.001$, $\eta_p^2\geq.43$) compared with experts who used low MTJ.

Conclusions: To provide health information online that is perceived as credible, experts should consider using similar language as the language used by the addressed audience. As it is often impossible to determine the exact makeup of an online audience, further research might investigate whether having experts explicitly declare which audience they intend to address can help people to more reliably assess an expert's trustworthiness. Furthermore, as people assess information differently depending on the context of online communication, it would be valuable for research to consider other aspects of the context beyond those of the audience.

(J Med Internet Res 2018;20(1):e30) doi:10.2196/jmir.8346

KEYWORDS

trust; health communication; social media; information seeking behavior; language

Introduction

Trust in Online Health Forums

For a member of the general public trying to find specific advice in an online forum, determining who to believe and on which information to rely is fraught with a number of challenges [1,2]. The main challenges are related to, on the one hand, the quality of information and the lack of quality control [3,4], and, on the other hand, the difficulty that laypersons have in adequately assessing, understanding, and judging the given information [2,5-7]. As people are increasingly using the Internet to obtain nutrition information [8,9], it is becoming more important to understand the associated challenges people face in accurately judging the information and the information providers.

To get correct and understandable health information, laypersons need to rely on qualified information providers, as their own expertise might lack in specific topics [10]. By relying on the information from an information provider, the layperson risks losing something or being disadvantaged in some way if the information is false [11-13]. Although people differ in their willingness to rely on others as a source of information and to judge the persuasive quality of a source (ie, epistemic trustworthiness) [14], how people determine if something is credible relates more to the objective features of the information itself [15]. However, judging credibility also involves subjective judgments of, for instance, a source's trustworthiness, expertise, and attractiveness [5,6]. Therefore, in this paper, trustworthiness and credibility are used as concepts dependent on each other, whereas epistemic trustworthiness focuses on the reliability of experts and credibility focuses on the reliability of the provided information.

Although finding health information online still poses risks to information seekers of being misinformed by unreliable information or information providers, having access to online information can be empowering. People want to find understandable health information online [7,16]. Moreover, if people can find understandable information and comprehend content-related issues, they may make better health-related decisions [2,17]. Accordingly, the most common reason people visit online health forums (eg, about nutrition) is to search for information [18,19].

Online health forums are visited by users from wide-ranging communities with various knowledge backgrounds (ie, medical professionals and medical novices, as well as patients and members of the general public). These forums allow users to exchange information, regardless of a user's knowledge of or expertise in a certain subject [18,20,21]. Thereby, the extent of professionals' engagement in forum discussions differs between forums [22], and the main user group is likely to be different in different forums. In some forums, professionals *discuss among themselves* [23,24], whereas in others, professionals advise a user group made up mainly of laypersons [25,26], and in others still, laypersons primarily exchange information with other laypersons [27-29].

However, in any online forum, the actual audience extends beyond the forum's designated (intended) audience; many users *lurk*, that is, they visit forums only passively to receive information rather than to post their own contribution [20,30]. Accordingly, several forms of expert-layperson communication may occur, such that different numbers of people with different levels of expertise are involved (eg, one expert is communicating with one layperson while simultaneously one expert is communicating with many others). Moreover, in online forums, these types of expert-layperson communications differ from offline expert-layperson communication. First, the health information provided in online forums reaches a broad layperson public [20], because online forums are not closed and hence can be accessed by a large public. Thus, even when an expert is communicating directly with a specific person in a forum, many other visitors to the forum (users) and, potentially, even the entire public could read the provided information. Second, the way in which experts use language becomes even more important in online forums because most often their written language is the only obvious cue that can be used by forum visitors to assess the expert's trustworthiness [31,32]. Regarding the used language in expert-layperson communication, it is particularly challenging when experts use technical jargon [33]. Using too much technical jargon could prevent the main goal to impart knowledge, as people with less background knowledge might not understand technical jargon [33,34] or find it less credible [35].

Although in online forums the forms of terminology vary from everyday language to highly technical jargon, users may expect a certain type of wording depending on the designated audience of forum types. Accordingly, users would need to recognize not only the specific language use of experts but also the forum context. Therefore, the study reported here addresses passive forum users' judgment of experts' trustworthiness and credibility based on the way the experts used language in different forum contexts. Online information seekers are at risk to interpret even accurate health information incorrectly (leading to misconceptions, misinformation, and harmful decisions) if they do not recognize the context cues that indicate a certain piece of health-related information as intended for specific audiences (also described as *context deficit* [2,4]). Similarly, the communication context plays a crucial role in the communication accommodation theory (CAT) [36], and it not only influences the communication itself (eg, the language style used) but also affects the perception of communication as well as the communication outcome.

Communication Accommodation Theory Online

Following the CAT [36,37], online health communication can be considered as an interplay between (1) the people communicating, (2) the language they are using, and (3) the context (ie, the immediate situation where communication takes place). For instance, the specific audience and the interlocutors' languages determine what variants of language are used [38]. People adjust their language using 2 key strategies: accommodation and nonaccommodation.

Accommodation relates to the process of using similar language relative to the perceived language of the audience, and it can be regarded as a tactic for conforming to the audience and emphasizing social belongingness [39]. Moreover, accommodation can lead to higher comprehension by the audience, as adapting one's language to that of the audience can occur after one perceives the audience's background knowledge [10,40]. Similarly, when an expert uses high amounts of medical technical jargon (MTJ) in a professional forum, as well as when they use low amounts of MTJ in an advisory forum, this is considered to be accommodation, as it refers to using similar language relative to the perceived language and background knowledge of the audience.

On the other hand, nonaccommodation relates to the process of using different language relative to the perceived language of the audience. In this context, nonaccommodation can be regarded as a tactic for showing that one has distinct values from the interlocutor [41]. Importantly though, nonaccommodative use of MTJ could inhibit comprehension and knowledge transfer [34]. Accordingly, when an expert uses high amounts of MTJ in an advisory forum or low amounts of MTJ in a professional forum, this can be considered nonaccommodation, as it refers to using different language relative to the perceived language of the audience.

Research on health communication often focuses on 2 essential communication styles used by doctors (ie, doctor-centered vs patient-centered communication) and their impact on patients' evaluation of doctors' and patients' behavioral changes [42,43]. Although doctor-centered communication is characterized by rational-cognitive proceedings without paying much attention to a patient's needs and feelings, patient-centered communication is more open, non-directive, and aims to actively engage patients by focusing on their psychological and social situation [43]. In patient-centered communication, doctors would tend to avoid using medical terms, as this takes patients' needs for understanding information and satisfaction with doctor-patient communication into account [44,45]. However, doctors often use more medical terms compared with even well-educated patients [46].

According to previous research on the effectiveness of communication styles, information providers that used a more doctor-centered communication style were perceived by participants to be less competent, less empathetic, and less trustworthy, and they led participants to report fewer changes in attitude compared with when information providers used a more patient-centered communication approach [43,47,48]. However, this does not mean that all patients equally prefer a patient-centered communication. Instead, a study on patients' preferred communication style showed that there are also 30.8% (77/250) patients who preferred a more doctor-centered communication style. Moreover, the remaining patients who preferred a more patient-centered communication style characterized their own doctor as patient-centered, appreciated when the doctor showed interest in the patient as a person, and were younger than 65 years [49]. Thus, doctors' communication styles should try to match patients' communication preferences to enhance patients' satisfaction and to promote patients' health care [48]. Accordingly, a match between a doctor's and a patient's communication style can be considered accommodation, and a mismatch can be considered nonaccommodation. For instance, when focusing on the wording, medical students answering an email inquiry were influenced by the technicality of inquiries and adapted their use of words to the level of technicality in the inquiry [50,51]. Similarly, users in online health forums adapted their language of replies to the language of the inquiry [52], and medical professionals of 7 major German health portals used no medical terms in reply to a question only when that question itself did not contain any medical terms [53]. In summary, these semiprofessionals and professionals used similar language styles, relative to those used in the health inquiries (ie, accommodation). In contrast, medical students in early semesters with high biomedical orientation (ie, having concepts of scientific and evidence-based medicine) replied more scientifically and less emotionally to all wordings of patients' queries and, hence, did not use wording similar to what the patients used (ie, nonaccommodation) [39].

This Study and Hypothesis

This study sheds light on whether passive forum users are sensitive to experts' use of accommodative or nonaccommodative language; we assessed this by varying the designated audience of an online forum and by varying the amount of MTJ (ie, the wording of information) experts used in their responses to forum inquiries. Specifically, we wanted to know how these 2 factors influenced users' judgments of the experts' trustworthiness and credibility. In face-to-face settings, people seem to think it is more appropriate for a speaker to accommodate their language to match that of the addressed audience compared with when the speaker does not accommodate [54]. In addition, accommodative language use is also strongly accompanied with people's perceived credibility [55]. Is this also true for information processing in online health forums?

Research on health communication suggests that professionals should consider information seekers' needs to promote information seekers' understanding and health [42,43,47-49]. Accordingly, professionals' communication styles can match the information seekers' communication styles and, therefore, reflect these needs [50-53]. Hence, the accommodative use of MTJ should increase each passive user's confidence that audiences are receiving appropriate information from competent advice givers. Thus, we hypothesized that accommodative language use would lead to higher credibility and trustworthiness ratings than nonaccommodative language use (ie, the interaction effect of the designated audience in forum types and amount of MTJ). We assumed that in a professional forum, a high amount of MTJ would lead to higher ratings for the dependent variables of credibility of information [56], trustworthiness of experts [14], and accommodation by experts [33] compared with when a high amount of MTJ was used in an advisory forum. Similarly, we assumed that in an advisory forum, a low amount of MTJ would lead to higher ratings for the dependent variables compared with when a low amount of MTJ was used in a professional forum.

As the aim of the investigation was to examine whether people judge the credibility of information and an expert's trustworthiness according to the expert's accommodative (or nonaccommodative) use of MTJ in 2 different types of forums, we had participants assess not only the expert's trustworthiness and credibility of information but also the perceived accommodation by the experts to the audience.

Methods

Design

We applied a 2 × 2 mixed design, with the independent factors being designated audience of the forum type (in a *professional forum*, the main audience is other medical experts who are exchanging information with each other, and in an *advisory forum*, the main audience is laypersons) and the amount of MTJ used (high vs low).

The designated audience was manipulated between conditions. Professional forums were introduced as forums that are mainly used by medical experts to exchange technical content and to take part in scientific discussions. In contrast, advisory forums were introduced as forums that are mainly used by patients and other nonprofessionals to inform themselves about health-related questions, to take part in explanation-oriented discussions, and to receive answers from medical professional (see Multimedia Appendix 1). Both introductions for forum types declared that the following posts were written by experts. Hence, the information providers in both forums had the same status as experts, as people judge trustworthiness and the credibility of information based on the information providers' expertise [57]. In addition to describing the forum type in the overall introduction, every post about a nutrition myth was introduced by a slightly different short description about the forum type (according to the condition) to promote participants' awareness of the forum context (Multimedia Appendix 2).

MTJ was varied within subjects; either the expert's post included language using Greek or Latin expressions or it included synonymous everyday German expressions (see Multimedia Appendix 3). In addition, 1 complex sentence from the version containing the high level of MTJ was split into 2 less complex sentences for the version having a low level of MTJ.

Procedure

Participants completed an online survey via the platform Questback EFS Survey and answered questions regarding demographic and control variables before they were randomly assigned to one of the forum type conditions (professional forum vs advisory forum). Depending on which condition they were assigned to, they either read an introduction explaining that the following posts came from a forum used mainly by medical professionals to discuss among themselves or they read an introduction explaining that the posts came from a forum in which medical professionals advice a user group mainly made up of laypersons. This overall introduction of forum type was presented for at least 16 seconds to ensure participants saw it. Every participant had to complete 10 survey sites, each of which included a short description of the specific forum type, a moderated question to introduce the nutrition topic, and a screenshot of the expert's post. Posts were presented in a fixed order to control the influence of different nutrition myths.

Materials

Nutrition myths were chosen as the content domain for this study because it is an interesting (hot) topic [58,59]. All nutrition myths (eg, coffee and dementia, olive oil and cardiac infarction, healthy number of eggs; 10 topics in total, see Multimedia Appendix 3) were based on our online searches for frequent and typical nutrition myths in online forums. To explain the science behind the myths, we used the textbook Nutrition Science by de Groot and Farhadi [60], and we formulated certain scientific limitations by using relativizations (eg, there is some evidence). To make these texts look like real experts' posts, one online forum was created; texts were posted; and the date, amount, and origin of posts were made unrecognizable (screenshot of a survey site showing an expert's post, see Multimedia Appendix 4). Overall, each participant received 10 experts' posts about a nutrition myth (5 high in amount of MTJ and 5 low in amount of MTJ). Posts did not differ in average length. In the German language, both forms of medical terminology exist: everyday medical terms in German as well as medical technical terms that stem from Greek or Latin (eg, blood pressure-lowering vs antihypertensive) [10,61]. MTJ not only contains medical technical terms that stem from Greek or Latin but also contains highly complex sentences [62]. Hence, the experts' posts at the high MTJ level included medical technical terms (Greek and Latin) and 1 complex sentence, whereas the experts' posts at the low MTJ level contained no medical technical terms (they were replaced with everyday medical synonyms) and no complex sentences. Synonymous usage of these terms was ensured by referring to a medical dictionary [63] (notice that in German, nearly every everyday medical term has a Greek or Latin synonym [64]).

Participants

An analysis of power (assuming 1 − ß=.80; Cohen f=.25) was used in advance to determine the sample size (n=92). In total, data from 106 participants were collected. Three participants were excluded from data analysis because at the end of the survey they specified that they did not want to provide their data for research purposes. Three additional participants were excluded because their time to complete the survey (ranging from 267 min to 870 min) took more than one standard deviation (SD) above the overall mean duration of all participants (mean 49.3 min, SD 90.6); we decided to exclude these participants because they were likely engaged in other activities when they were answering the survey. Furthermore, 2 more participants were excluded because they reported professional prior knowledge—1 studied health promotion and 1 was a trained paramedic.

Hence, in sum, data from 98 participants (72 females) aged 18 to 54 years (mean 25.18, SD 6.33) were analyzed. A total of 92 participants indicated German as their first language (the other 6 participants had been speaking German for an average of 11.17 years [SD 5.31]). Moreover, 93 participants specified that they were currently studying or had studied at the university level and hence had a university-entrance diploma. On average, they reported studying for 4.77 semesters (SD 3.04).

Furthermore, 52 majored in psychology, and the others came from various disciplines (19 from law, economics, and social science; 8 from natural science, engineering science, and mathematics; 8 from linguistic and cultural studies; 5 from teacher training; and 1 from sports science).

Participants reported that they used a computer for an average of 29.06 (SD 20.02) hours per week and the Internet for an average of 26.71 (SD 22.75) hours per week. A total of 81 out of 98 participants (83%) reported that they used text messaging (short message service, SMS) daily. Regarding email use, 39 out of 98 participants (40%) reported using it daily; 49 participants (50%) at least several times per week. In addition, most participants did not regularly visit forums: 72 out of 98 participants (74%) stated they used a forum less than once per month. Furthermore, participants reported that their prior knowledge of and motivation to be informed about the topic *health and nutrition* (from 1=I agree to 5=I disagree) to be on average 2.33 (SD 0.77); the 4 items they scored themselves on were as follows: I deal a lot with the topic *health and nutrition*; To me, it is important to eat healthy; I read a lot about *health and nutrition* on the Internet; and I am familiar with the topic *health and nutrition*.

Participants were invited by a link on several online platforms run by a big German university and a German remote university. They were automatically excluded from the survey if they had stated that they were currently studying or had studied medicine or nutrition science. We excluded them because these students' high amount of prior knowledge may have influenced their judgments of credibility and trust in the information within the forum posts. A total of 47 participants were assigned to the professional forum condition and 51 were assigned to the advisory forum condition; they received a 10€ voucher as reimbursement. In addition, the survey access was automatically denied when participants attempted to use a mobile phone; thus, it allowed us to control the screen size of the devices used. The average duration of completed surveys was 37.99 min (SD 15.70).

Dependent Measures

This study employed 3 dependent variables: credibility of information, trustworthiness of experts, and perceived accommodation by experts.

Credibility of Information

Participants were asked to indicate whether they agreed with the given information (How much would you agree with the answer?) and to judge the credibility of the information using 4 additional items adopted from a measurement scale assessing trust in journalism [56]. We further added the item *I would like to ask someone else*. Overall, these 6 items yielded an internal consistency of Cronbach alpha=.86. Participants rated items on a 5-point Likert scale from 1 (I strongly disagree) to 5 (I strongly agree).

Trustworthiness of Experts

Trustworthiness of experts was assessed by Muenster Epistemic Trustworthiness Inventory (METI) [14]. METI was designed to measure the epistemic trustworthiness of unknown sources and, therefore, is useful in online settings, because online information about the source is often scarce. METI consists of 14 items and is composed of 3 subscales. The subscale Expertise reflects the participant's perception of the expert as truly knowledgeable, intelligent, and highly trained in her domain. The subscale Integrity is related to the expert's good character and values, and reflects the participant's perception of the expert as a person who is acting in line with norms. The Benevolence subscale represents the participant's perception of whether the expert acts in accordance with the interest of others.

To rate the trustworthiness of each expert, participants used 7-point scales (eg, 1=competent to 7=incompetent) containing acronym adjectives that represented the subscale to which the items belonged. The item unselfish-selfish (not included in the revised version of METI) was added. Internal consistencies were Cronbach alpha=.95 for the subscale Expertise (6 items), Cronbach alpha=.89 for Integrity (5 items), and Cronbach alpha=.87 for Benevolence (4 items). Overall, all 15 items yield a Cronbach alpha=.95 (METI Score).

Perceived Accommodation by Experts

The perceived accommodation by experts was assessed using an adaption of the recipient orientation scale (ROS) [33], which refers to how experts are perceived to adjust their language to an audience. The dimension Audience Design reflects the perceived willingness of the expert to adapt to the audience (eg, the expert can imagine how it is to know little about this topic). The dimension Subjective Comprehension is related to the participant's self-reported understanding about the topic at hand (eg, I understood the content). In addition, the dimension Emotional Evaluation is related to the participant's self-reported liking of reading the expert's posts (eg, it is exciting to read the response of the expert). Overall, all 13 items yielded a Cronbach alpha=.90. Internal consistencies were Cronbach alpha=.81 for Audience Design (6 Items), Cronbach alpha=.72 for Subjective Comprehension (3 Items), and Cronbach alpha=.73 for Emotional Evaluation (4 Items).

Results

Preliminary Analysis

In a preliminary analysis, we assessed the differences between subsamples of each forum condition for the expected control variables. A multivariate analysis revealed no differences in groups regarding any of the participant demographic variables, their reported usage of forums, the Internet or computers, or the 4 items that assessed self-reported general prior knowledge, all $F_{1,91}<3.78$, $P>.06$. Moreover, a second preliminary multivariate analysis included the within factor MTJ and revealed no differences between the forum conditions regarding the 3 control items that assessed subjective familiarity (I'm familiar with the content of the response), subjective complexity (This is a complex issue), and interest (The response is interesting) for each topic. Hence, none of these items were included as a control variable in the following analysis.

An alpha level of .05 was set for a multivariate variance analysis with repeated measures, where MTJ was the within subject factor and designated audience was the between subject factor.

All tests were one sided. For each participant, we used averaged values of the dependent variables for all the 5 high MTJ and all the 5 low MTJ responses; hence, for each participant, 2 average values—one for the low MTJ condition and one for the high MTJ condition—were used for further analysis.

Credibility of Information

There were no main effects of MTJ ($F_{1,96}$=0.07, P=.40) and of designated audience ($F_{1,96}$=0.52, P=.24) regarding credibility. However, the multivariate analysis yielded a significant interaction of MTJ and designated audience ($F_{1,95}$=3.10, P=.04, η_p^2=.031), with higher credibility judgments given for the condition that used high amounts of MTJ in professional forums (mean 3.05, SD 0.38) compared with the condition of using low amounts of MTJ in professional forums (mean 2.97, SD 0.47). Conversely, in advisory forums, low amounts of MTJ (mean 3.00, SD 0.60) led to higher credibility judgments than high amounts of MTJ (mean 2.90, SD 0.44). Figure 1 illustrates the interaction of MTJ and designated audience in terms of credibility of information.

Trustworthiness of Experts

Both main effects showed significance: MTJ affected all the 3 subscales of the METI; Expertise $F_{1,96}$=37.54, P<.001, η_p^2=.28, Integrity, $F_{1,96}$=10.77, P=.001, η_p^2=.101, and Benevolence, $F_{1,96}$=9.75, P=.002, η_p^2=.092. High amounts of MTJ led to higher expertise ratings (mean 2.82, SD 0.80) than did low amounts of MTJ (mean 3.19, SD 0.79). However, low amounts of MTJ led to higher Integrity (mean 3.42, SD 0.62) and Benevolence ratings (mean 3.47, SD 0.67) compared with conditions of high amounts of MTJ (Integrity: mean 3.57, SD 0.55; Benevolence: mean 3.66, SD 0.57).

The designated audience affected the METI score $F_{1,96}$=3.54, P=.03, η_p^2=.036, and the subscales of Integrity, $F_{1,96}$=3.54, P=.03, η_p^2=.036, and Benevolence, $F_{1,96}$=6.11, P=.001, η_p^2=.06. Thereby, experts in professional forums were judged as being more trustworthy (mean 3.21, SD 0.63), of greater integrity (mean 3.39, SD 0.63), and more benevolent (mean 3.43, SD 0.65) than experts in advisory forums (METI score: mean 3.40, SD 0.57; Integrity: mean 3.59, SD 0.53; Benevolence: mean 3.70, SD 0.56). There was no significant interaction of MTJ and designated audience regarding trustworthiness, all $F_{1,95}$<1.84, P>.09.

Perceived Accommodation by Experts

Significant main effects of MTJ and designated audience were found for perceived accommodation by experts. MTJ affected the overall score of the ROS and all subscales, all $F_{1,96}$≥72.17, P<.001, η_p^2≥.43, where more expert accommodation was found in the low MTJ condition. The analysis also revealed a significant main effect of designated audience on the subscale Emotional Evaluation, $F_{1,96}$=3.35, P=.04, η_p^2=.03, where participants ascribed emotional evaluation to experts in professional forums (mean 3.10, SD 0.51) than to experts in advisory forums (mean 2.92, SD 0.58). There were no significant interaction effects of MTJ and designated audience regarding perceived accommodation by experts: all $F_{1,95}$<2.61, P>.06.

Table 1 shows descriptive values for the factors MTJ and designated audience for each of the dependent variables. Table 2 shows values for the multivariate analysis of variance, including the within factor of MTJ and the between factor of designated audience for each of the dependent variables.

Figure 1. Interaction of medical technical jargon and designated audience in terms of credibility of information (1=strongly disagree to 5=strongly agree).

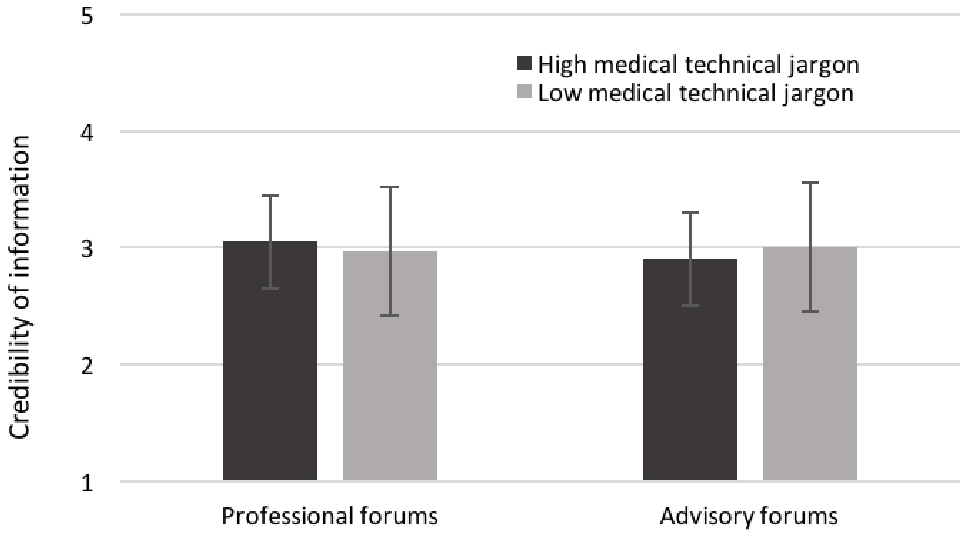


Table 1. Descriptive statistics.

Factors	Professional forum, mean (SD)	Advisory forum, mean (SD)	Total, mean (SD)
Trustworthiness of experts			
METI [a] score			
High MTJ[b]	3.16 (0.63)	3.42 (0.50)	3.30 (0.58)
Low MTJ	3.26 (0.63)	3.41 (0.64)	3.34 (0.63)
Total	3.21 (0.63)	3.42 (0.57)	3.31 (0.60)
Expertise			
High MTJ	2.71 (0.81)	2.92 (0.78)	2.82 (0.80)
Low MTJ	3.12 (0.79)	3.26 (0.79)	3.19 (0.79)
Total	2.91 (0.80)	3.11 (0.79)	3.01 (0.79)
Integrity			
High MTJ	3.44 (0.62)	3.70 (0.46)	3.58 (0.55)
Low MTJ	3.34 (0.63)	3.49 (0.60)	3.42 (0.62)
Total	3.39 (0.63)	3.59 (0.53)	3.49 (0.58)
Benevolence			
High MTJ	3.48 (0.63)	3.83 (0.45)	3.66 (0.57)
Low MTJ	3.38 (0.68)	3.56 (0.67)	3.47 (0.67)
Total	3.43 (0.65)	3.70 (0.56)	3.56 (0.61)
Credibility of information			
High MTJ	3.05 (0.38)	2.90 (0.44)	2.97 (0.42)
Low MTJ	2.97 (0.47)	3.00 (0.60)	2.99 (0.54)
Total	3.01 (0.42)	2.95 (0.52)	2.98 (0.47)
Perceived accommodation by experts			
ROS [c] score			
High MTJ	2.79 (0.42)	2.62 (0.49)	2.70 (0.46)
Low MTJ	3.51 (0.46)	3.48 (0.44)	3.49 (0.45)
Total	3.15 (0.44)	3.05 (0.47)	3.10 (0.45)
Audience design			
High MTJ	2.58 (0.46)	2.43 (0.53)	2.51 (0.49)
Low MTJ	3.43 (0.45)	3.49 (0.42)	3.46 (0.43)
Total	3.01 (0.45)	2.96 (0.48)	2.98 (0.47)
Subjective comprehension			
High MTJ	3.25 (0.64)	3.10 (0.64)	3.17 (0.64)
Low MTJ	3.80 (0.55)	3.70 (0.55)	3.75 (0.55)
Total	3.52 (0.60)	3.40 (0.60)	5.16 (0.60)
Emotional evaluation			
High MTJ	2.76 (0.49)	2.54 (0.57)	2.64 (0.52)
Low MTJ	3.41 (0.55)	3.30 (0.60)	3.35 (0.58)
Total	3.10 (0.52)	2.92 (0.58)	6.00 (0.55)

[a] METI: Muenster Epistemic Trustworthiness Inventory.

[b] MTJ: medical technical jargon.

[c] ROS: recipient orientation scale.

Table 2. Multivariate analysis of variance including the factors medical technical jargon (within subjects) and designated audience (between subjects) for each of the dependent variables.

Factors and tests[a]	Degrees of freedom	*F*	*P* [a]	η_p^2
MTJ [b] (within)				
Trustworthiness of experts				
METI[c] score	1	0.951	.17	0.010
Expertise	1	37.543	<.001	0.281
Integrity	1	10.772	<.001	0.101
Benevolence	1	9.751	.001	0.092
Credibility	1	0.068	.40	0.001
Perceived accommodation by experts				
ROS[d] score	1	178.611	<.001	0.650
Audience design	1	219.130	<.001	0.695
Subjective comprehension	1	72.172	<.001	0.429
Emotional evaluation	1	11.492	<.001	0.537
Forum type (between)				
Trustworthiness of experts				
METI score	1	3.537	.03	0.036
Expertise	1	1.469	.11	0.015
Integrity	1	3.537	.03	0.036
Benevolence	1	6.113	.01	0.060
Credibility	1	0.523	.24	0.005
Perceived accommodation by experts				
ROS score	1	2.047	.08	0.021
Audience design	1	0.486	.24	0.005
Subjective comprehension	1	1.541	.11	0.016
Emotional evaluation	1	3.346	.04	0.034
MTJ forum type				
Trustworthiness of experts				
METI score	1	1.425	.12	0.015
Expertise	1	0.405	.26	0.004
Integrity	1	1.575	.11	0.016
Benevolence	1	1.841	.09	0.019
Credibility	1	3.101	.04	0.031
Perceived accommodation by experts				
ROS score	1	1.497	.11	0.015
Audience design	1	2.609	.06	0.026
Subjective comprehension	1	0.178	.34	0.002
Emotional evaluation	1	0.741	.20	0.008

[a] All tests were one-sided.

[b] MTJ: medical technical jargon.

[c] METI: Muenster Epistemic Trustworthiness Inventory.

[d] ROS: recipient orientation scale.

Discussion

Principal Findings

The designated audience and technicality of medical jargon, in general, impacted peoples' assessment of trustworthiness. An expert who used a low level of MTJ was perceived to be less competent but of more integrity, more benevolent, and more accommodative in terms of addressing the audience than an expert who used a high level of medical jargon. In addition, experts in the context of a professional forum were judged to be more trustworthy—respectively, to have more integrity and benevolence—than experts in advisory forums. Furthermore, information was judged to be more credible when medical jargon was accommodative rather than nonaccommodative to the designated audience. Participants judged information written with high amounts of MTJ to be more credible in professional forums, whereas they judged the same information to be less credible in advisory forums. At the same time, participants judged information written with low amounts of MTJ to be more credible in advisory forums, whereas they judged the same information to be less credible in professional forums. Hence, our hypothesis that accommodative language use should lead to higher perceived information credibility, greater expert trustworthiness, and greater perceived accommodation by the expert compared with nonaccommodative language use can be confirmed in terms of information credibility.

Limitations

Regarding the representation of the sample, it is striking that most participants did not use forums very often. Although unexperienced forum users clearly took into account where information was presented and whether the use of medical jargon was appropriate in a given context, it is unclear whether their unfamiliarity with forum use affected participants' abilities to adequately assess trustworthiness or credibility. However, information literacy influences perceptions of credibility [6], and more experienced Internet users consider media in general to be more credible [65,66]. Thus, these findings suggest that participants who are more experienced in using forums might have used other factors to apply their judgments than those used by the participants in our sample—regardless of whether their judgments would have been adequate. Therefore, further research is needed to illuminate whether participants who use forums more often would find similar effects.

In addition, the significant interaction in terms of credibility produced only a small effect size. However, research on psychological media effects faces several challenges (as do many other research fields) that usually lead to small effect sizes. These challenges include measuring media exposure in a reliable and valid way, while at the same time considering certain conditional variables [67].

Surprisingly, the supposed appropriateness of forum type and language use did not affect the perceived accommodation by experts. Instead, due to our operationalization of accommodation by experts, participants may have judged an expert's accommodation as it fits not only to the designated forum audience but also to an undefined group of over-hearers and eavesdroppers, including the participants themselves. However, we assume that participants did not judge each expert's accommodation in terms of accommodating the specific person who inquired, because due to our reformulation of inquiry it is unlikely that participants expected the expert to address a specific user's post. Even so, participants might have faced challenges in identifying whether the experts intended to address the designated audiences or the entire public, as online forums are not privately closed [20,68].

Implications

In terms of trustworthiness, our results indicate a few guidelines for experts who are interested in providing information in online health forums. Results show that, generally, providing information in professional forums appears to be more trustworthy. Moreover, using high amounts of MTJ, generally, makes the expert appear to be more competent but also to have less integrity and be less benevolent. Furthermore, experts should aim to use accommodative language and hence use MTJ more appropriately, as taking the audience into account increased people's perceived credibility of information provided by experts and could promote information seekers' understanding of information [10,40,50]. However, further research should be done on whether having experts explicitly declare which audience they intend to address—either a specific online audience or the entire public—can help people to more reliably and accurately assess the expert's trustworthiness and the credibility of their information. Moreover, future research might include aspects of biased information processing, such as in instances where forum users have more well-defined prior attitudes toward the presented information, which users likely did not have in this study [69]. In terms of the relatively high unexplained statistical variance, which resulted in a small effect size for the credibility effect, future research could further consider the influence of individuals' differences in or underlying mechanisms of information processing [67].

Especially online, people face several challenges in identifying factors that indicate who should be trusted and on which information to rely. Factors that can be used to assess the credibility of information are, for instance, the author's credentials, their expertise, and if they use comprehensible language [6,7,35]. In addition, people should also use the context of online communication to make these judgments [2,4]. However, to our knowledge, only few studies have investigated whether users are sensitive to context cues when judging the credibility of online information [70]. Future research, therefore, should not only investigate whether people are sensitive to different online contexts but should also consider cues about online context other than the designated audience.

Conclusions

The results of our study clearly illustrate that users keep track of wording and the context of information when reading online health information. It is interesting to see that not only did both experimental factors impact the assessment, but so did the interaction of the 2 factors. According to methodology, this study illustrates how specific research questions can be addressed by varying central features of online forums in an experimental study. Thus, this study is able to offer more specific insights on trustworthiness and credibility assessment in online settings than research usually provides [6,71]. By focusing on 2 relevant aspects of online forums (ie, the context, namely, designated audience, and the MTJ), this study helps to assess the effect of appropriateness between language use and the context of online communication. A future challenge will be to specify the appropriateness of language use not only in terms of online audience but also in terms of other aspects of online communication contexts.

Acknowledgments

The authors thank Corinna Sperling for her help in implementing the experimental setting, data collection, and processing. The authors thank Celeste Brennecka for language editing. The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this paper: this research was supported by the German Research Foundation (DFG) within the framework of Research Training Group 1712/2: Trust and Communication in a Digitized World. The German Research Foundation had no involvement in study design, data collection, analysis and interpretation, and the decision to submit the paper for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Framing of designated audience of forum types (translation from German).

[PDF File (Adobe PDF File), 39KB - jmir_v20i1e30_app1.pdf]

Multimedia Appendix 2

Short forum descriptions (translation from German).

[PDF File (Adobe PDF File), 42KB - jmir_v20i1e30_app2.pdf]

Multimedia Appendix 3

High vs low medical technical jargon for each of the 10 nutrition myths and the introductory question for expert's posts (translation from German).

[PDF File (Adobe PDF File), 110KB - jmir_v20i1e30_app3.pdf]

Multimedia Appendix 4

Screenshot of a survey site showing an expert's post.

[PDF File (Adobe PDF File), 405KB - jmir_v20i1e30_app4.pdf]

References

1. Sbaffi L, Rowley J. Trust and credibility in web-based health information: a review and agenda for future researcha. J Med Internet Res 2017 Jun 19;19(6):e218 [FREE Full text] [doi: 10.2196/jmir.7579] [Medline: 28630033]
2. Eysenbach G. Credibility of health information digital media: New perspectives implications for youth. In: Metzger MJ, Flanagin AJ, editors. Digital media, youth, credibility. Cambridge, MA: The MIT Press; 2008:123-154.
3. Modave F, Shokar NK, Peñaranda E, Nguyen N. Analysis of the accuracy of weight loss information search engine results on the internet. Am J Public Health 2014 Oct;104(10):1971-1978. [doi: 10.2105/AJPH.2014.302070] [Medline: 25122030]
4. Eysenbach G, Diepgen TL. Towards quality management of medical information on the internet: evaluation, labelling, and filtering of information. Br Med J 1998 Nov 28;317(7171):1496-1500 [FREE Full text] [Medline: 9831581]
5. Metzger MJ, Flanagin AJ. Credibility and trust of information in online environments: the use of cognitive heuristics. J Pragmat 2013 Dec;59:210-220. [doi: 10.1016/j.pragma.2013.07.012]
6. Choi W, Stvilia B. Web credibility assessment: conceptualization, operationalization, variability, and models. J Assn Inf Sci Tec 2015 May 13;66(12):2399-2414. [doi: 10.1002/asi.23543]
7. Sillence E, Briggs P, Harris PR, Fishwick L. How do patients evaluate and make use of online health information? Soc Sci Med 2007 May;64(9):1853-1862. [doi: 10.1016/j.socscimed.2007.01.012] [Medline: 17328998]
8. Pollard CM, Pulker CE, Meng X, Kerr DA, Scott JA. Who uses the Internet as a source of nutrition and dietary information? An Australian population perspective. J Med Internet Res 2015 Aug 26;17(8):e209 [FREE Full text] [doi: 10.2196/jmir.4548] [Medline: 26310192]

9. Fox S. Pew Internet. 2011. Health topics URL: http://www.pewinternet.org/files/old-media//Files/Reports/2011/PIP_Health_Topics.pdf [accessed 2015-03-10] [WebCite Cache ID 6WuhdBEU9]
10. Bromme R, Jucks R. Discourse and expertise: The challenge of mutual understanding between experts and laypeople. In: Schober M, Rapp DN, Britt MA, editors. The Routledge Handbook of Discourse Processes. New York, NY: Taylor & Francis; 2017:222-246.
11. Mayer RC, Davis JH, Schoorman FD. An integrative model of organizational trust. Acad Manage Rev 1995 Jul 01;20(3):709-734. [doi: 10.5465/AMR.1995.9508080335]
12. Friedman B, Khan PH, Howe DC. Trust online. Commun ACM 2000;43(12):34-40. [doi: 10.1145/355112.355120]
13. Kelton K, Fleischmann K, Wallace W. Trust in digital information. J Am Soc Inf Sci 2008 Feb 01;59(3):363-374. [doi: 10.1002/asi.v59:3]
14. Hendriks F, Kienhues D, Bromme R. Measuring laypeople's trust in experts in a digital age: the muenster epistemic trustworthiness inventory (METI). PLoS One 2015;10(10):e0139309 [FREE Full text] [doi: 10.1371/journal.pone.0139309] [Medline: 26474078]
15. Freeman KS, Spyridakis JH. An examination of factors that affect the credibility of online health information. Tech Commun 2004 May;51(2):239-263 [FREE Full text]
16. Chiu Y. Probing, impelling, but not offending doctors: the role of the internet as an information source for patients' interactions with doctors. Qual Health Res 2011 Dec;21(12):1658-1666. [doi: 10.1177/1049732311417455] [Medline: 21799204]
17. Eysenbach G. The impact of the Internet on cancer outcomes. CA Cancer J Clin 2003;53(6):356-371. [Medline: 15224975]
18. Hajli MN, Sims J, Featherman M, Love PE. Credibility of information in online communities. JSM 2014 May 22;23(3):238-253. [doi: 10.1080/0965254X.2014.920904]
19. Kim S, Yoon J. The use of an online forum for health information by married Korean women in the United States. Inf Res 2012;17(2):1-18 [FREE Full text]
20. Anesa P, Fage-Butler A. Popularizing biomedical information on an online health forum. Iberica 2015;29:105-128.
21. Vennik FD, Adams SA, Faber MJ, Putters K. Expert and experiential knowledge in the same place: patients' experiences with online communities connecting patients and health professionals. Patient Educ Couns 2014 May;95(2):265-270. [doi: 10.1016/j.pec.2014.02.003] [Medline: 24598314]
22. Johnsen JK, Rosenvinge JH, Gammon D. Online group interaction and mental health: an analysis of three online discussion forums. Scand J Psychol 2002 Dec;43(5):445-449. [Medline: 12500784]
23. Reddit. URL: https://www.reddit.com/ [accessed 2017-07-04] [WebCite Cache ID 6rhlN6Vqd]
24. Doctors.net.uk. URL: https://forum.doctors.net.uk/ [accessed 2017-07-04] [WebCite Cache ID 6rhlV7hYX]
25. MedHelp. URL: http://medhelp.org [accessed 2017-07-04] [WebCite Cache ID 6rhleSLK1]
26. Doctors.net.uk. URL: https://forums.netdoctor.co.uk/ [accessed 2017-07-04] [WebCite Cache ID 6rhljrq7A]
27. Inspire. URL: https://www.inspire.com [accessed 2017-09-13] [WebCite Cache ID 6tRmeRwpX]
28. Dailystrength. URL: https://www.dailystrength.org/ [accessed 2017-09-13] [WebCite Cache ID 6tRmzMpMa]
29. MacLean D, Gupta S, Lembke A, Manning C, Heer J. An analysis of an online Health forum dedicated to addiction recovery. Forum77; 2015 Mar 14 Presented at: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing; March 14-18, 2015; Vancouver, BC, Canada p. 14-18. [doi: 10.1145/2675133.2675146]
30. Nonnecke B, Preece J. Why lurkers lurk. 2001 Dec Presented at: Americas Conference on Information Systems; 2001; Boston, MA p. 1-10 URL: http://aisel.aisnet.org/amcis2001/294
31. Jucks R, Linnemann G, Thon F, Zimmermann M. Trust the words. Insights into the role of language in trust building in a digitalized world. In: Blöbaum B, editor. Trust and Communication in a Digitized World. Cham: Springer International Publishing; 2016:225-237.
32. Liao W, Bazarova NN, Yuan YC. Expertise judgment and communication accommodation in linguistic styles in computer-mediated and face-to-face groups. Communic Res 2016 Feb 05:-. [doi: 10.1177/0093650215626974]
33. Bromme R, Jucks R, Runde A. Barriers and Biases in Computer-Mediated Knowledge Communication. Boston, MA: Springer; 2005:98-118.
34. Jucks R, Paus E. What makes a word difficult? Insights into the mental representation of technical terms. Metacogn Learn 2011 Nov 17;7(2):91-111. [doi: 10.1007/s11409-011-9084-6]
35. Thon FM, Jucks R. Believing in expertise: How authors' credentials and language use influence the credibility of online health information. Health Commun 2017 Jul;32(7):828-836. [doi: 10.1080/10410236.2016.1172296] [Medline: 27466693]
36. Giles H. Communication accommodation theory: Negotiating personal and social identities in context. Cambridge, UK: Cambridge University Press; 2016.
37. Coupland J, Coupland N, Giles H, Henwood K. Accommodating the elderly: invoking and extending a theory. Lang Soc 2008 Dec 18;17(01):1 [FREE Full text] [doi: 10.1017/S0047404500012574]
38. Bell A. Language style as audience design. Lang Soc 2008 Dec 18;13(02):145-204 [FREE Full text] [doi: 10.1017/S004740450001037X]
39. Bientzle M, Griewatz J, Kimmerle J, Küppers J, Cress U, Lammerding-Koeppel M. Impact of scientific versus emotional wording of patient questions on doctor-patient communication in an internet forum: a randomized controlled experiment with medical students. J Med Internet Res 2015;17(11):e268 [FREE Full text] [doi: 10.2196/jmir.4597] [Medline: 26607233]

40. Gasiorek J, Giles H, Soliz J. Accommodating new vistas. Lang Commun 2015 Mar;41:1-5. [doi: 10.1016/j.langcom.2014.10.001]
41. Giles H, Ogay T. Communication Accommodation Theory. In: Whaley BB, Samter W, editors. Explaining Communication: Contemporary Theories and Exemplars. Mahwah, NJ: Lawrence Erlbaum; 2007:293-310.
42. Stevenson FA, Cox K, Britten N, Dundar Y. A systematic review of the research on communication between patients and health care professionals about medicines: the consequences for concordance. Health Expect 2004 Sep;7(3):235-245. [doi: 10.1111/j.1369-7625.2004.00281.x] [Medline: 15327462]
43. Bientzle M, Fissler T, Cress U, Kimmerle J. The impact of physicians' communication styles on evaluation of physicians and information processing: a randomized study with simulated video consultations on contraception with an intrauterine device. Health Expect 2017 Oct;20(5):845-851. [doi: 10.1111/hex.12521] [Medline: 27860037]
44. Kripalani S, Jacobson TA, Mugalla IC, Cawthon CR, Niesner KJ, Vaccarino V. Health literacy and the quality of physician-patient communication during hospitalization. J Hosp Med 2010;5(5):269-275 [FREE Full text] [doi: 10.1002/jhm.667] [Medline: 20533572]
45. Sim D, Yuan S, Yun J. Health literacy and physician-patient communication: a review of the literature. Int J Commun Health 2016;10:101-114.
46. Koch-Weser S, Rudd RE, Dejong W. Quantifying word use to study health literacy in doctor-patient communication. J Health Commun 2010 Sep;15(6):590-602 [FREE Full text] [doi: 10.1080/10810730.2010.499592] [Medline: 20812122]
47. Saha S, Beach MC. The impact of patient-centered communication on patients' decision making and evaluations of physicians: a randomized study using video vignettes. Patient Educ Couns 2011 Sep;84(3):386-392 [FREE Full text] [doi: 10.1016/j.pec.2011.04.023] [Medline: 21600723]
48. Pollard CM, Howat PA, Pratt IS, Boushey CJ, Delp EJ, Kerr DA. Preferred tone of nutrition text messages for young adults: focus group testing. JMIR Mhealth Uhealth 2016 Jan 19;4(1):e1 [FREE Full text] [doi: 10.2196/mhealth.4764] [Medline: 26787115]
49. Swenson SL, Buell S, Zettler P, White M, Ruston DC, Lo B. Patient-centered communication: do patients really prefer it? J Gen Intern Med 2004 Nov;19(11):1069-1079 [FREE Full text] [doi: 10.1111/j.1525-1497.2004.30384.x] [Medline: 15566435]
50. Jucks R, Paus E, Bromme R. Patients' medical knowledge and health counseling: what kind of information helps to make communication patient-centered? Patient Educ Couns 2012 Aug;88(2):177-183. [doi: 10.1016/j.pec.2012.01.011] [Medline: 22365590]
51. Bromme R, Jucks R, Wagner T. How to refer to ‘diabetes’? Language in online health advice. Appl Cognit Psychol 2005 Jul;19(5):569-586. [doi: 10.1002/acp.1099]
52. Kimmerle J, Bientzle M, Cress U. “Scientific evidence is very important for me”: the impact of behavioral intention and the wording of user inquiries on replies and recommendations in a health-related online forum. Comput Human Behav 2017 Aug;73:320-327. [doi: 10.1016/j.chb.2017.03.064]
53. Jucks R, Bromme R. Choice of words in doctor-patient communication: an analysis of health-related internet sites. Health Commun 2007;21(3):267-277. [doi: 10.1080/10410230701307865] [Medline: 17567258]
54. Brown BL, Giles H, Thakerar JN. Speaker evaluations as a function of speech rate, accent and context. Lang Commun 1985 Jan;5(3):207-220. [doi: 10.1016/0271-5309(85)90011-4]
55. Soliz J, Giles H. Relational and identity processes in communication: a contextual and meta-analytical review of Communication Accommodation Theory. Annals of the International Communication Association 2016 May 18;38(1):107-144. [doi: 10.1080/23808985.2014.11679160]
56. Matthes J, Kohring M. Operationalization of trust in journalism [in German]. M&K 2003;51(1):5-23. [doi: 10.5771/1615-634x-2003-1-5]
57. Yang Q, Beatty M. A meta-analytic review of health information credibility: belief in physicians or belief in peers? HIM J 2016 Aug;45(2):80-89. [doi: 10.1177/1833358316639432] [Medline: 27105476]
58. Wangberg S, Andreassen H, Kummervold P, Wynn R, Sørensen T. Use of the internet for health purposes: trends in Norway 2000-2010. Scand J Caring Sci 2009 Dec;23(4):691-696. [doi: 10.1111/j.1471-6712.2008.00662.x] [Medline: 19473317]
59. Trends.google. URL: https://trends.google.de/trends/explore?q=nutrition [accessed 2017-09-14] [WebCite Cache ID 6tSITon7q]
60. de Groot H, Farhadi J, editors. Ernährungswissenschaft. Haan-Gruiten: Europa-Lehrmittel; 2015.
61. Wulff HR. The language of medicine. J R Soc Med 2004 Apr;97(4):187-188 [FREE Full text] [Medline: 15056745]
62. Langer I, Schulz von Thun F, Tausch R, editors. Sich verständlich ausdrücken. München: Ernst Reinhard Verlag; 2011.
63. Reuter P, editor. Springer Taschenwörterbuch Medizin. Berlin: Springer Verlag; 2005.
64. Jucks R, Becker B, Bromme R. Lexical entrainment in written discourse? Is expert's word use adapted to the addressee? Discourse Process 2008 Dec 11;45(6):497-518. [doi: 10.1080/01638530802356547]
65. Flanagin AJ, Metzger MJ. Perceptions of internet information credibility. Journal Mass Commun Q 2016 Jun 25;77(3):515-540. [doi: 10.1177/107769900007700304]

66. Zulman DM, Kirch M, Zheng K, An LC. Trust in the internet as a health resource among older adults: analysis of data from a nationally representative survey. J Med Internet Res 2011;13(1):e19 [FREE Full text] [doi: 10.2196/jmir.1552] [Medline: 21324832]
67. Valkenburg P, Peter J. Five challenges for the future of media-effects research. Int J Commun 2013;7:197-215 [FREE Full text]
68. Bernstein M, Bashy E, Burke M, Karrer B. Quantifying the invisible audience in social networks. 2013 May 02 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; 2013; Paris, France p. 21-30. [doi: 10.1145/2470654.2470658]
69. Nauroth P, Gollwitzer M, Bender J, Rothmund T. Social identity threat motivates science-discrediting online comments. PLoS One 2015;10(2):e0117476 [FREE Full text] [doi: 10.1371/journal.pone.0117476] [Medline: 25646725]
70. Hu Y, Shyam Sundar S. Effects of online health sources on credibility and behavioral intentions. Communic Res 2009 Nov 25;37(1):105-132. [doi: 10.1177/0093650209351512]
71. Sundar S. The MAIN Model: A heuristic approach to understanding technology effects on credibility. In: Metzger MJ, Flanagin AJ, editors. Digital Media, Youth, and Credibility. Cambridge, MA: The MIT Press; 2008:73-100.

Abbreviations

CAT: communication accommodation theory
METI: Muenster Epistemic Trustworthiness Inventory
MTJ: medical technical jargon
ROS: recipient orientation scale
SMS: short message service

Edited by G Eysenbach; submitted 05.07.17; peer-reviewed by P Schulz, J Kimmerle; comments to author 24.08.17; revised version received 27.09.17; accepted 29.10.17; published 23.01.18.

Please cite as:
Zimmermann M, Jucks R
How Experts' Use of Medical Technical Jargon in Different Types of Online Health Forums Affects Perceived Information Credibility: Randomized Experiment With Laypersons
J Med Internet Res 2018;20(1):e30
URL: http://www.jmir.org/2018/1/e30/
doi:10.2196/jmir.8346
PMID:29362212

Review

Digital Clinical Communication for Families and Caregivers of Children or Young People With Short- or Long-Term Conditions: Rapid Review

Xavier Armoiry[1], PharmD, PhD; Jackie Sturt[2], PhD; Emma Elizabeth Phelps[1], PhD; Clare-Louise Walker[1], MBChB; Rachel Court[1], MA; Frances Taggart[1], PhD; Paul Sutcliffe[1], PhD; Frances Griffiths[1,3], MBBS, PhD; Helen Atherton[1], PhD

[1]Warwick Medical School, University of Warwick, Coventry, United Kingdom

[2]Florence Nightingale Faculty of Nursing, Midwifery & Palliative Care, Kings College London, London, United Kingdom

[3]Centre for Health Policy, University of the Witwatersrand, Johannesburg, South Africa

Corresponding Author:
Frances Griffiths, MBBS, PhD
Warwick Medical School
University of Warwick
Gibbet Hill Road
Coventry, CV4 7AL
United Kingdom
Phone: 44 02476522534
Email: f.e.griffiths@warwick.ac.uk

Abstract

Background: The communication relationship between parents of children or young people with health conditions and health professionals is an important part of treatment, but it is unclear how far the use of digital clinical communication tools may affect this relationship.

Objective: The objective of our study was to describe, assess the feasibility of, and explore the impact of digital clinical communication between families or caregivers and health professionals.

Methods: We searched the literature using 5 electronic databases. We considered all types of study design published in the English language from January 2009 to August 2015. The population of interest included families and caregivers of children and young people aged less than 26 years with any type of health condition. The intervention was any technology permitting 2-way communication.

Results: We included 31 articles. The main designs were randomized controlled trials (RCTs; n=10), cross-sectional studies (n=9), pre- and postintervention uncontrolled (pre/post) studies (n=7), and qualitative interview studies (n=2); 6 had mixed-methods designs. In the majority of cases, we considered the quality rating to be fair. Many different types of health condition were represented. A breadth of digital communication tools were included: videoconferencing or videoconsultation (n=14), and Web messaging or emails (n=12). Health care professionals were mainly therapists or cognitive behavioral therapists (n=10), physicians (n=8), and nurses (n=6). Studies were very heterogeneous in terms of outcomes. Interventions were mainly evaluated using satisfaction or acceptance, or outcomes relating to feasibility. Clinical outcomes were rarely used. The RCTs showed that digital clinical communication had no impact in comparison with standard care. Uncontrolled pre/post studies showed good rates of satisfaction or acceptance. Some economic studies suggested that digital clinical communication may save costs.

Conclusions: This rapid review showed an emerging body of literature on the use of digital clinical communication to improve families' and caregivers' involvement in the health management of children or young people. Further research with appropriate study designs and longer-term outcome measures should be encouraged.

Trial Registration: PROSPERO CRD42016035467; http://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD 42016 035467(Archived by WebCite at http://www.webcitation.org/6vpgZU1FU)

(J Med Internet Res 2018;20(1):e5) doi:10.2196/jmir.7999

KEYWORDS

digital clinical communication; professional-family relations; family; caregivers; young adult; children; child health

Introduction

Digital clinical communication can be defined as a means of communication between a clinician and a person, when the clinician or the person (or both) is (or could be) mobile when sending or receiving the communication, in a 2-way, synchronous or asynchronous manner, and for clinical care purposes only [1]. The use of digital clinical communication technologies has been extensively described due to their capacity to facilitate communication between health care professionals and patients [2-4]. With children and young people being prolific users of these technologies, there is much speculation about the potential feasibility of digital communication between children and young people with health conditions, on the one hand, and their health professionals, on the other, as a way to meet the specific needs of this population [5]. Indeed, young people are particularly at risk of disengaging from health services and experience poorer health outcomes [1,6-8]. Although the technology is promising, the effectiveness of digital communication with patients or parents of children and young people in health care on outcomes has not been clearly demonstrated [9].

Parental involvement and parent-health professional relationships are an important part of the treatment journey of children and young people with health conditions, but it is unclear what impact digital communication has on these relationships, particularly as young people transition into using adult services. Digital communication with health care providers may also be used by families involved in the management of health conditions in much younger pediatric populations, where parents are fully acting as communicator with health services [10,11].

Given the wide spectrum of these digital tools and the different modalities used by families, we aimed to review the literature and the emerging conceptualization of the topic.

For this purpose, we chose a rapid review method, which can be defined as a form of knowledge synthesis in which components of the systematic review process are simplified or omitted to produce information in a timely manner [12]. Rapid reviews are useful in fields where change is ongoing [13], such as in the development of digital technologies.

The aim of this review was to establish the current evidence base for the use of digital clinical communication for families and caregivers of children or young people with short- or long-term conditions.

Our objectives were to describe existing digital communication use by health professionals with families or caregivers of children or young persons with short- or long-term conditions, to assess the feasibility of using these technologies, and to explore their impact on (1) family and caregivers' outcomes, (2) children and young people's outcomes, (3) health professionals' outcomes, and (4) health service delivery and health economics outcomes.

Methods

We report this review according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [14].

Inclusion and Exclusion Criteria

Our inclusion criteria were (1) studies of any design published in the English language (except conference abstracts and articles with fewer than 5 participants); (2) family members or caregivers of individuals less than 26 years of age and presenting with all types of health condition; we extended the age range to 25 years to include literature on the issue of transition from pediatric to adult health services; and (3) all forms of 2-way digital communication between family members or caregivers and health professionals, including email; social networking sites; mobile telephony; short message service (SMS) text messaging systems; video- and teleconferencing; online forums; and electronic monitoring.

We excluded (1) technologies involving only 1-way data transmission, communication between family members or caregivers and children or young people, and communication between family members or caregivers; and (2) technologies involving 2-way communication between children or young people and professionals if there was no involvement of family members or caregivers.

In accordance with the Cochrane Consumers & Communication Group's taxonomy [15], we were interested in studies that included patient outcomes, family and caregiver outcomes, health professional outcomes, and health service delivery and economic outcomes. We assessed feasibility according to the simple definition "the state or degree of being easily or conveniently done." This was assessed via patient and health care professional outcomes and via reporting of technical or usability concerns. We assessed the impact of digital clinical communication only through controlled studies, while we assessed feasibility using both controlled and uncontrolled studies.

Search Strategy

We developed a literature search strategy to search 5 electronic databases (MEDLINE [through Ovid], Embase [Ovid], MEDLINE In-Process & Other Non-Indexed Citations [Ovid], PsycINFO [ProQuest], and Cochrane Library [Wiley]) in August 2015 for relevant literature published in or after January 2009. We chose this time period to include the most recent digital communication tools. We used a combination of free-text and thesaurus terms for the concepts of technology, clinical communication, population, and families and caregivers" to identify related literature (Multimedia Appendix 1). We also searched citations and the reference lists of relevant studies. Furthermore, we hand searched within JMIR journals (themes: *Clinical Communication*; *Electronic Consultation and Telehealth*; *Email & Web-Based Communication*; *Personal*

Health Records; *Patient-Accessible Electronic Health Records*; *Patient Portals*) over the same period.

Screening and Analysis

Two independent reviewers screened all identified bibliographic records by title and abstract. We obtained full-text articles for all remaining records, and these were read by 1 reviewer. The final list of included studies at full-text level was validated by 2 other reviewers. One reviewer extracted data from all the included studies using an a priori-defined, prepiloted extraction sheet that was designed by the same reviewer and included data on the population, intervention, comparator (where relevant), and outcomes. A second independent reviewer double-checked the extracted data. A third reviewer resolved any disagreement.

We assessed the quality of randomized controlled trials (RCTs), economic evaluations, and qualitative research articles using the Critical Appraisal Skills Programme checklists [16]. For pre- and postintervention uncontrolled (pre/post) studies, cross-sectional and observational studies, and non-RCTs, we assessed study quality using checklists published by the US National Institutes of Health [17]. We used these checklists because, in the context of a rapid review, these provide a useful way to assess several different study designs consistently and quickly. For mixed-methods studies, we undertook a quality assessment for each study method. Quality was assessed independently by 2 reviewers. Any disagreement between reviewers was resolved by consensus or with recourse to a third reviewer. We rated the overall quality of studies as poor, fair, or good. This rating was assessed by the 2 reviewers through discussion, accounting for each study's limitations as emphasized by the items within the checklists. For example, a study with high risk of bias or major flaws based on several checklist items translated to a rating of poor quality. Conversely, a study with low risk of bias or free from major flaw translated to a rating of good quality. We summarized study, intervention, population, and outcome characteristics narratively and in summary tables. Meta-analysis and statistical pooling were not possible owing to the heterogeneity of interventions and health conditions that we identified. This rapid review was registered in PROSPERO (CRD42016035467).

Results

Characteristics of Included Studies

Of the 1156 identified records (including 10 additional records from reference lists of relevant studies), we removed 956 not meeting our inclusion criteria at title and abstract stage, leaving 200 articles to be examined at full-text review. Among these, there was 1 systematic review of interest [18], which we checked for the presence of potentially relevant articles. We excluded 169 articles not meeting our inclusion criteria, leading to a total of 31 included publications (Figure 1).

The main study designs were RCTs (n=10), cross-sectional studies (n=9), pre/post studies (n=7), and qualitative interview studies (n=2) (Multimedia Appendix 2) [10,11,19-47]. Of the studies, 6 had mixed-methods designs. Most of the studies were conducted in the United States (n=17), while the other main locations were Australia (n=3), the Netherlands (n=3), and Sweden (n=3).

We identified a broad range of conditions: traumatic brain injury (n=5), the management of prematurity and associated consequences (n=3), atopic dermatitis (n=2), autism spectrum disorder (n=2), type 1 diabetes (n=2), palliative care for different types of diseases (n=2), and anorexia nervosa (n=1). The range of the mean age of children and young people was 24 days to 20.4 years. Of the 31 selected articles, 23 included a mainly pediatric population (age <12 years), while 7 mainly included adolescents (12-18 years) and 1 mainly included young adults (>18 years). Of the 7 articles including adolescents, the involvement of families and caregivers in using the digital clinical communication technology was the key component of the intervention because of the health condition of adolescents (mainly traumatic brain injury or cerebral disability).

We found no study on the use of digital communication by families or caregivers of young people during the transition to adult care.

We rated the majority of studies as being of fair quality. The main limitation was the uncertainty as to whether the participants were representative of those in the general population (for studies where generalizability was relevant). Multimedia Appendix 3 reports the full description of included studies and our quality assessment of them.

Description of Existing Digital Communication

A range of digital communication channels were used across studies. Videoconferencing allowing consultation was predominantly used (n=14) followed by emails or Web messaging systems (n=12) (Table 1).

Health care professionals were mainly therapists or cognitive behavioral therapists (n=10), physicians (n=8), and nurses (n=6). Stand-alone interventions were used in 13 studies, while in 18 studies the digital communication was included within a wider intervention such as a Web-based therapy or Web-based system (portal, telemedicine, telehealth) that allowed 2-way communication between health care professionals and families.

Family and Caregiver Outcomes

Multimedia Appendix 4 summarizes the results and Multimedia Appendix 5 reports the complete results. The benefits reported include removing barriers to communication [47], providing reassurance for those with chronic illness [47], and feeling supported for those adolescents with eating disorders [29]. The majority of families felt satisfied with the digital communication tools [40,41]. All the families found these tools easy to use [20,30,35,44] and some said they would recommend them [39,40].

We found no studies reporting a difference in family and caregiver outcomes between the group using digital communication and the control or alternative intervention group (eg, telephone) [10,21,26,28,36,43].

Figure 1. Flow diagram of study identification and analysis for inclusion. CBT: cognitive behavioral therapy.

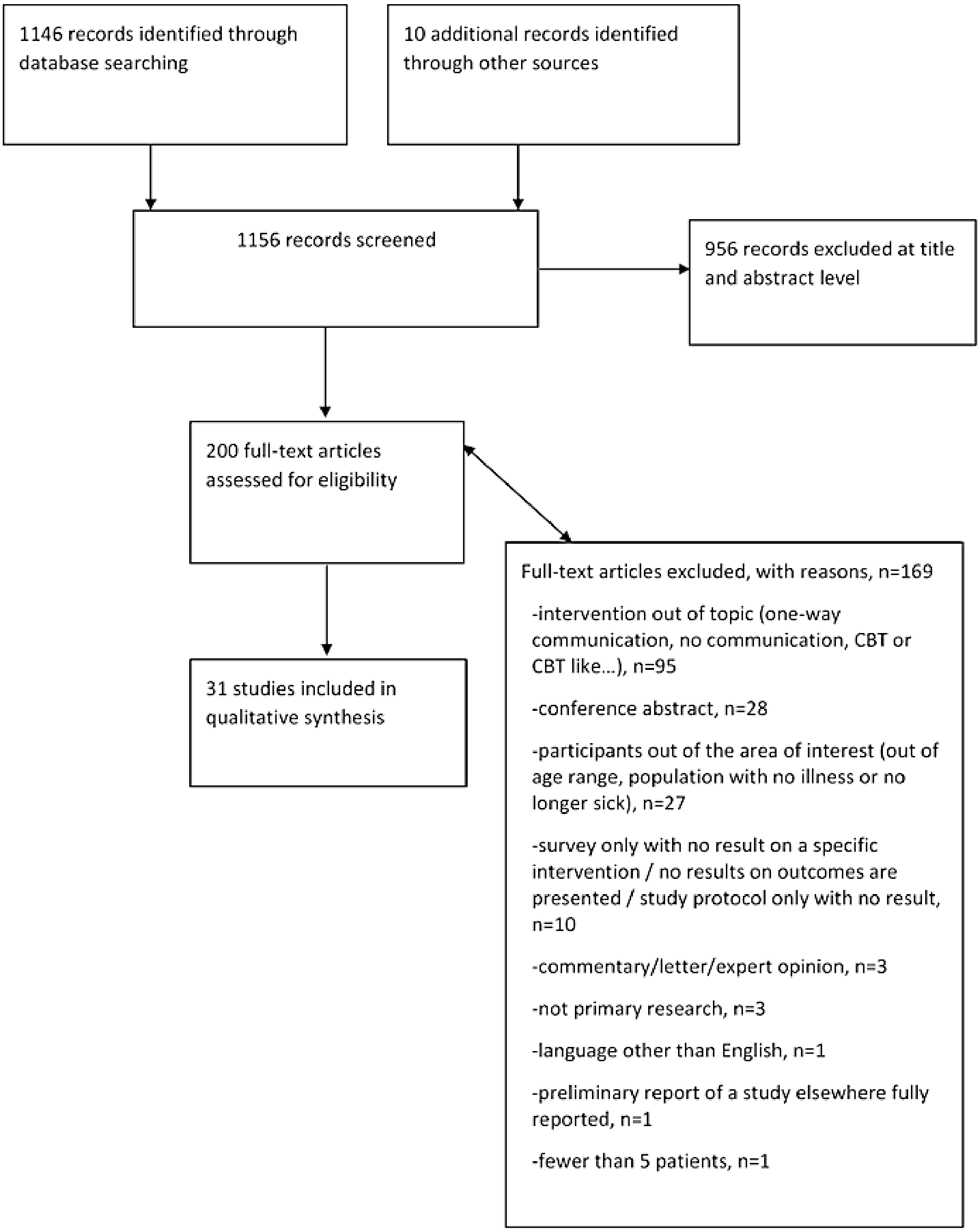


Table 1. Description of digital clinical communication tools as identified in the rapid review.

Digital clinical communication tool	Studies (n)
Videoconferencing systems allowing consultations	14
Web messaging or emails	12
Web chat	2
Web-based telemedicine systems with no other element	2
Short message service (SMS) text messaging	1

A total of 6 studies reported improvements for families and caregivers after the intervention compared with baseline [10,26,29,33,34,36], while 2 reported no difference [30,32]. For the parents of adolescents with eating disorders, there was an improvement on the Eating Disorders Symptom Impact Scale after participation in Web chat sessions with fellow parents and a clinical psychologist [29].

In an intervention providing early autism training to parents, the Maternal Behavior Rating Scale (designed to assess the quality of maternal interactive behavior with children with learning difficulties) score was increased compared with baseline after interventions that included videoconferencing and a telehealth-delivered curriculum [34] or Web-based learning modules [33]. Caregiver distress was reduced after the intervention [26]; the delivery of pediatric palliative home care via videoconsultation improved scores on health-related quality of life [36]. Lastly, the number of skin care treatments given by parents of children with atopic dermatitis was increased after Web-based consultations [10]. Of these 6 studies [10,26,29,33,34,36], 3 [10,26,36] found similar improvements from baseline in the control or alternative intervention group.

In an intervention delivering care for children with obesity and comparing face-to-face delivery with telehealth, parents rated telehealth consultations lower than face-to-face consultations when asked whether the provider explained things about the child's health in a way that was easy to understand [43].

Patient Outcomes

Of the 6 studies reporting patients' clinical outcomes (Multimedia Appendix 4), 4 [10,28,33,34] found significant improvements (change compared with baseline, $P<.05$) after the intervention. Vismara et al [34] found that after speech therapy the rate of child vocalizations and their joint attention increased, while another study [33] found that child social communication behaviors improved following the use of an intervention comprising videoconferencing and learning modules. Among the 2 studies evaluating stand-alone digital communication using patients' clinical outcomes [10,28], none showed a difference compared with the control group. In e-counselling for behavioral problems, Becker [28] found the frequency and severity of the child's disruptive behaviors reduced at 2 weeks compared with baseline, but this difference was found in both the digital communication group and the telephone group. Similarly, in children with atopic dermatitis, the Objective Severity Scoring of Atopic Dermatitis scores were improved at 12 months in both the intervention (remote dermatology consultations via Web messaging) and control groups [10].

Digital communication was described as helpful by patients with traumatic brain injury [45] and as easy to use by the majority of patients using a portal for diabetes management [44].

Health Professional Outcomes

In the 2 studies reporting health care professionals' perceptions [20,30], their response was positive overall. In 1 study testing the feasibility of Skype and FaceTime updates with parents in the neonatal intensive care unit, 94% of providers rated the ease of using videoconferencing as excellent or good and more than 90% perceived videoconferencing to be reliable [30]. Similarly, in the home health care of premature infants, most nurses were motivated to use the information and communication technology; however, some were reluctant [20], feeling, for example, that the use of digital communication by the families should be discouraged in general, since these activities took families' attention away from their infant.

Health Service Delivery and Economic Outcomes

A total of 4 studies evaluated the effect of digital communication on health service delivery outcomes [10,37,38,46]. Of these, 2 reported the frequency of use, with 1 study of a Web-based tool for atopic dermatitis finding 8.3 messages sent per participant over 12 months [10], and 1 study of an SMS text messaging tool for caregivers of patients with disabilities finding 6.25 messages sent per participant over 3 months [46]. In the latter study, the content of digital communication was also described, with participants using SMS text messaging for social interaction and to ask questions [46]. Email was used predominantly to provide participants with information about common diseases and treatments. In an online portal for patients with chronic conditions, 64% of participants used the portal instead of calling their health care provider on at least one occasion [37,38].

In the Australian context, pediatric palliative care by videoconsultation at home saved costs versus face-to-face consultation during hospital-based consultations or during home visits [36].

There were 2 studies on parental management of children with atopic dermatitis evaluating economic outcomes [10,25]. One compared Web-based consultations versus a control group where participants were encouraged to seek treatment through traditional means [10]; except for hospital admissions, there were fewer health care visits (general practitioner visits, outpatient consultations, emergency visits, and complementary therapists visits) at 12 months in both groups, but there was no significant difference between the 2 groups. The other study reported that an eHealth portal was saving €594 per patient in the first year, mainly through a reduction in work absenteeism, the probability of eHealth reducing costs being 73% or greater after sensitivity analyses [25].

Technical Problems

In 4 studies [29,30,34,44], participants reported experiencing technical problems. Vismara et al [34], who used telehealth for early autism training, found that all participants experienced some degree of frustration when using the videoconferencing program, including the audio or webcam not working or the Internet connection freezing. Passwords were reported as a barrier to using a diabetes Web portal and Web messaging [44], with the procedure for replacing lost passwords and creating one's own password found to create problems. Using a videoconferencing system in the neonatal intensive care unit, some parents experienced technical problems, such as frozen screens, attributed to poor Internet connections at the parent's home [30]. Finally, in a study focused on adolescent eating disorders [29], 3 out of 13 parents experienced technical problems, the nature of which was not described, during the Web chat sessions with a therapist.

Discussion

Principal Findings

The majority of studies were quantitative research, which explored the impact of digital communication in a broad range of conditions. Videoconsultations were often used. Most articles reported experiences of families of young children where the communication was with a parent. Findings of these studies of the use of digital clinical communication were equivocal, with no clear benefits in relation to patient, caregiver, or health care professional outcomes reported, but no adverse events reported either. Where digital clinical communication had been explored qualitatively, the key themes were the perception of the removal of barriers to communication and the formation of networks for communication facilitated by these technologies. Digital communications were found to be acceptable in most of the studies. Overall, parents were satisfied with their experience and perceived benefit.

Findings for economic benefit were equivocal: 2 studies found economic benefits of digital communication in comparison with face-to-face consultations, while 1 showed a neutral impact on resource use.

Our rapid review also had a specific focus on the use of digital communications by families and caregivers of young people in the transition period between pediatric and adult services. We had extended the age range for children and young people in this review to 26 years to include this period. Transition care is particularly challenging for the young person, the parent or caregiver, and the health care team [48]. Digital communication affords opportunities to young people to communicate privately with their health care team as they transition toward assuming greater responsibility for their health from their parents. Within the scope of the transition period, a greater emphasis could be placed on using digital technologies to enhance communication between health care professionals and patients directly, rather than through their parents. However, guidance indicates that the involvement of parents, combined with that of young people, is important at the transition stage [49,50]. This is why we were also interested in the use of digital communication between health professionals and parents during the transition period. Our rapid review showed that this population was absent from the identified literature, indicating a lack of evidence relating to the impact of family and caregiver involvement within this scope.

We are aware of a systematic review that assessed telehealth tools and interventions to support family caregivers of pediatric, adult, and older patients with chronic diseases [18]. This review of 64 articles published over 1997-2014 concluded that telehealth can positively affect care provided by family caregivers. Our work differs from the previous review in that we were interested only in the impact of digital clinical communication for families or caregivers of children or young people with health conditions; this excluded adults and older patients.

Other comprehensive reviews have investigated the impact of communication technologies on young people with mental health conditions [5] or with diabetes [9], and concluded that the benefit of using these technologies was unclear or inconclusive. However, these 2 reviews had no specific focus on family or caregiver involvement, which further justifies our work.

Limitations and Strengths

Our rapid review presents several limitations. This work being a rapid review and not a systematic review [13], we chose not to undertake a gray literature search aimed at identifying unpublished studies, some of which may have failed to demonstrate the feasibility of these new technologies. Thus, the absence of publication bias in this area cannot be excluded. Overall, as we shortened time frames for literature searching and article retrieval, we may have missed some relevant information, which generates less certainty in our conclusions compared with using traditional systematic reviews [51]. As previously stated, we used here a rapid review method because the purpose of this type of review is to aid emergent decisions in health care settings.

As part of the rapid review process, we used simple checklists to assess study quality consistently and quickly. The use of detailed checklists may have enabled a more thorough quality assessment. However, as we were not excluding studies according to quality, we used these checklists in the context of a rapid review to give the reader an overview of the quality of the studies we included.

Many studies [19,21,31,38,44,47,52] examined interventions comprising more than one component, making it difficult to determine whether the results were due to the use of digital communication or other aspects of the intervention. For the purposes of this rapid review designed to scope the field using a narrative synthesis, we included these multifaceted studies, but future reviewers may wish to specifically limit their analysis to those studies where effects of the intervention can be separately assessed. In addition, we acknowledge that incorporating a wide range of outcomes does mean that some interpretation might be missed in narratively summarizing these, and focusing on particular outcome groupings of interest would allow for more in-depth synthesis.

As the included studies came from only 9 different countries, the generalizability of the findings beyond these settings is limited. Several of the studies came from the same groups of authors, working in specific clinical areas. These studies may therefore be indicative of clinical enthusiasts reporting their work. This has been observed in other reviews of digital interventions [3].

Although pre/post studies were not helpful to measure the direct impact of digital communication in the study populations, these studies were informative to address our objective to assess the feasibility of these technologies—that is, to verify their capacity to work correctly and be usable by health care professionals, families, and caregivers.

Regarding the assessment of feasibility, we emphasized in our quality assessment that a limitation of the studies was the uncertainty on whether the participants in RCTs and pre/post studies were representative of those in the general population, which again raises some issues of generalizability for these data. Indeed, the finding that some digital communication tools were usable by the included families and caregivers does not mean that the entire population would be able to use these with the same benefit. However, the use of digital communication technologies is becoming increasingly ubiquitous in our societies.

There is an emerging body of literature on the use of digital communication to improve families' and caregivers' involvement in the management of children and young people. Overall, these interventions show promise, but the evidence base to support them is lacking since, overall, we rated just 10 studies as good quality. Based on results from RCTs showing no differences between digital communication technologies and standard care, some authors suggest that such means of communication could be used to reach populations in underserved areas within large territories such the United States or Australia.

However, we would recommend further confirmatory studies using RCT designs to be conducted with a mid- to long-term perspective and using outcomes other than feasibility outcomes only. An RCT can be designed as mainly a noninferiority or superiority study. If the digital clinical communication is aimed at replacing an existing means of communication between health professional and families or caregivers, we believe a noninferiority design may be most appropriate. With such a design, the choice of the noninferiority margin, which corresponds to some loss of efficacy that might be accepted, could be easily justified by accounting for other benefits that digital clinical communication might have over standard care. If the use of digital clinical communication is added to standard care in one particular health condition in order to fulfill an unmet need, we believe a superiority design should be implemented. In this case, we would also recommend an economic evaluation to be conducted alongside so as to explore the economic impact the new intervention may have.

Given the absence of studies, future research could also be conducted with a specific focus on the use of digital communications in the transition period between pediatric and adult services given the special needs of this population [53,54].

Conclusion

This rapid review showed an emerging body of literature suggesting the feasibility and acceptance of, and good rates of satisfaction with, digital technologies to enhance communication between health care professionals and the families or caregivers of children or young people under their care. However, we found no clear impact of these technologies on outcomes. Further evaluations based on comparative studies with larger sample sizes are needed to confirm these preliminary results and should investigate the impact of digital communication in terms of quality and organization of care, as well as the associated economic outcomes. An important topic for future research could be the evaluation of digital communication involving parents and caregivers in the management of young people during the transition from pediatric to adult health services.

Acknowledgments

This rapid review was part of the LYNC study funded by the National Institute of Health Research, UK (call: 12/209 HS&DR Researcher Led).

Conflicts of Interest

CLW, FG, and HA are funded for other research by the National Institute of Health Research, UK.

Multimedia Appendix 1

Record of searches.

[PDF File (Adobe PDF File), 105KB - jmir_v20i1e5_app1.pdf]

Multimedia Appendix 2

Characteristics of included studies.

[PDF File (Adobe PDF File), 44KB - jmir_v20i1e5_app2.pdf]

Multimedia Appendix 3

Characteristics of included studies and quality assessment.

[PDF File (Adobe PDF File), 52KB - jmir_v20i1e5_app3.pdf]

Multimedia Appendix 4

Summary of results.

[PDF File (Adobe PDF File), 61KB - jmir_v20i1e5_app4.pdf]

Multimedia Appendix 5

Complete results.

[PDF File (Adobe PDF File), 59KB - jmir_v20i1e5_app5.pdf]

References

1. Griffiths F, Bryce C, Cave J, Dritsaki M, Fraser J, Hamilton K, et al. Timely digital patient-clinician communication in specialist clinical services for young people: a mixed-methods study (The LYNC Study). J Med Internet Res 2017 Apr 10;19(4):e102 [FREE Full text] [doi: 10.2196/jmir.7154] [Medline: 28396301]
2. Ye J, Rust G, Fry-Johnson Y, Strothers H. E-mail in patient-provider communication: a systematic review. Patient Educ Couns 2010 Aug;80(2):266-273 [FREE Full text] [doi: 10.1016/j.pec.2009.09.038] [Medline: 19914022]
3. Atherton H, Sawmynaden P, Sheikh A, Majeed A, Car J. Email for clinical communication between patients/caregivers and healthcare professionals. Cochrane Database Syst Rev 2012;11:CD007978. [doi: 10.1002/14651858.CD007978.pub2] [Medline: 23152249]
4. Newhouse N, Lupiáñez-Villanueva F, Codagnone C, Atherton H. Patient use of email for health care communication purposes across 14 European countries: an analysis of users according to demographic and health-related factors. J Med Internet Res 2015;17(3):e58 [FREE Full text] [doi: 10.2196/jmir.3700] [Medline: 25798912]
5. Martin S, Sutcliffe P, Griffiths F, Sturt J, Powell J, Adams A, et al. Effectiveness and impact of networked communication interventions in young people with mental health conditions: a systematic review. Patient Educ Couns 2011 Nov;85(2):e108-e119. [doi: 10.1016/j.pec.2010.11.014] [Medline: 21239133]
6. Health and Social Care Information Centre. National Diabetes Audit 2010-2011. Report 1: Care Processes and Treatment Targets. Leeds, UK: Health and Social Care Information Centre; 2012. URL: https://digital.nhs.uk/media/20306/National-Diabetes-Audit-2010-2011-Report1-Care-processes-and-treatment-targets-V4/Any/nati-diab-audi-10-11-care-proc-rep-V4 [accessed 2017-12-21] [WebCite Cache ID 6vsgPnVrK]
7. Singh SP. Transition of care from child to adult mental health services: the great divide. Curr Opin Psychiatry 2009 Jul;22(4):386-390. [doi: 10.1097/YCO.0b013e32832c9221] [Medline: 19417667]
8. Heery E, Sheehan AM, While AE, Coyne I. Experiences and outcomes of transition from pediatric to adult health care services for young people with congenital heart disease: a systematic review. Congenit Heart Dis 2015;10(5):413-427. [doi: 10.1111/chd.12251] [Medline: 25659600]
9. Sutcliffe P, Martin S, Sturt J, Powell J, Griffiths F, Adams A, et al. Systematic review of communication technologies to promote access and engagement of young people with diabetes into healthcare. BMC Endocr Disord 2011;11:1 [FREE Full text] [doi: 10.1186/1472-6823-11-1] [Medline: 21210964]
10. Bergmo TS, Wangberg SC, Schopf TR, Solvoll T. Web-based consultations for parents of children with atopic dermatitis: results of a randomized controlled trial. Acta Paediatr 2009 Feb;98(2):316-320. [doi: 10.1111/j.1651-2227.2008.01033.x] [Medline: 18795905]
11. Bradford N, Young J, Armfield NR, Bensink ME, Pedersen L, Herbert A, et al. A pilot study of the effectiveness of home teleconsultations in paediatric palliative care. J Telemed Telecare 2012 Dec;18(8):438-442. [doi: 10.1258/jtt.2012.gth103] [Medline: 23155114]
12. Tricco AC, Antony J, Zarin W, Strifler L, Ghassemi M, Ivory J, et al. A scoping review of rapid review methods. BMC Med 2015 Sep 16;13:224 [FREE Full text] [doi: 10.1186/s12916-015-0465-6] [Medline: 26377409]
13. Tsertsvadze A, Chen Y, Moher D, Sutcliffe P, McCarthy N. How to conduct systematic reviews more expeditiously? Syst Rev 2015 Nov 12;4:160 [FREE Full text] [doi: 10.1186/s13643-015-0147-7] [Medline: 26563648]
14. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol 2009 Oct;62(10):1006-1012. [doi: 10.1016/j.jclinepi.2009.06.005] [Medline: 19631508]
15. Cochrane Consumers and Communication. Outcomes of interest to the Cochrane Consumers & Communication Group. 2012 Jan. URL: http://cccrg.cochrane.org/sites/cccrg.cochrane.org/files/public/uploads/Outcomes.pdf [accessed 2017-05-09] [WebCite Cache ID 6qKUMHRpv]
16. CASP UK. Critical Appraisal Skills Programme (CASP): making sense of evidence. Oxford, UK: CASP UK; 2013. URL: http://www.casp-uk.net/ [accessed 2017-05-09] [WebCite Cache ID 6qKLsfYHc]
17. US Department of Health and Human Services, National Institutes of Health. Study quality assessment tools. 2016 Dec 01. URL: https://www.nhlbi.nih.gov/health-pro/guidelines/in-develop/cardiovascular-risk-reduction/tools [accessed 2017-05-09] [WebCite Cache ID 6qKT4z1qO]
18. Chi N, Demiris G. A systematic review of telehealth tools and interventions to support family caregivers. J Telemed Telecare 2015 Jan;21(1):37-44 [FREE Full text] [doi: 10.1177/1357633X14562734] [Medline: 25475220]
19. Grover M, Naumann U, Mohammad-Dar L, Glennon D, Ringwood S, Eisler I, et al. A randomized controlled trial of an Internet-based cognitive-behavioural skills package for carers of people with anorexia nervosa. Psychol Med 2011 Dec;41(12):2581-2591. [doi: 10.1017/S0033291711000766] [Medline: 21733215]

20. Gund A, Sjöqvist BA, Wigert H, Hentz E, Lindecrantz K, Bry K. A randomized controlled study about the use of eHealth in the home health care of premature infants. BMC Med Inform Decis Mak 2013 Feb 09;13:22 [FREE Full text] [doi: 10.1186/1472-6947-13-22] [Medline: 23394465]
21. Hanberger L, Ludvigsson J, Nordfeldt S. Use of a web 2.0 portal to improve education and communication in young patients with families: randomized controlled trial. J Med Internet Res 2013;15(8):e175 [FREE Full text] [doi: 10.2196/jmir.2425] [Medline: 23973555]
22. Looman WS, Antolick M, Cady RG, Lunos SA, Garwick AE, Finkelstein SM. Effects of a telehealth care coordination intervention on perceptions of health care by caregivers of children with medical complexity: a randomized controlled trial. J Pediatr Health Care 2015;29(4):352-363 [FREE Full text] [doi: 10.1016/j.pedhc.2015.01.007] [Medline: 25747391]
23. Petranovich CL, Wade SL, Taylor HG, Cassedy A, Stancin T, Kirkwood MW, et al. Long-term caregiver mental health outcomes following a predominately online intervention for adolescents with complicated mild to severe traumatic brain injury. J Pediatr Psychol 2015 Aug;40(7):680-688 [FREE Full text] [doi: 10.1093/jpepsy/jsv001] [Medline: 25682211]
24. Scharer K, Colon E, Moneyham L, Hussey J, Tavakoli A, Shugart M. A comparison of two types of social support for mothers of mentally ill children. J Child Adolesc Psychiatr Nurs 2009 May;22(2):86-98 [FREE Full text] [doi: 10.1111/j.1744-6171.2009.00177.x] [Medline: 19490279]
25. van Os-Medendorp H, Koffijberg H, Eland-de Kok PCM, van der Zalm A, de Bruin-Weller MS, Pasmans SGMA, et al. E-health in caring for patients with atopic dermatitis: a randomized controlled cost-effectiveness study of internet-guided monitoring and online self-management training. Br J Dermatol 2012 May;166(5):1060-1068. [doi: 10.1111/j.1365-2133.2012.10829.x] [Medline: 22268960]
26. Wade SL, Karver CL, Taylor HG, Cassedy A, Stancin T, Kirkwood MW, et al. Counselor-assisted problem solving improves caregiver efficacy following adolescent brain injury. Rehabil Psychol 2014 Feb;59(1):1-9 [FREE Full text] [doi: 10.1037/a0034911] [Medline: 24611923]
27. Wade SL, Walz NC, Carey J, McMullen KM, Cass J, Mark E, et al. A randomized trial of teen online problem solving: efficacy in improving caregiver outcomes after brain injury. Health Psychol 2012 Nov;31(6):767-776. [doi: 10.1037/a0028440] [Medline: 22746261]
28. Becker T. Effectiveness of Parent Call-In Versus E-Counseling Services in Treating Pediatric Behavior Problems Uncovered in a Primary Care Medical Encounter [doctoral thesis]. Newberg, OR: George Fox University; 2012. URL: http://digitalcommons.georgefox.edu/cgi/viewcontent.cgi?article=1100&context=psyd
29. Binford HRB, Le Grange D, Moessner M, Bauer S. Internet-based chat support groups for parents in family-based treatment for adolescent eating disorders: a pilot study. Eur Eat Disord Rev 2013 May;21(3):215-223. [doi: 10.1002/erv.2196] [Medline: 22949134]
30. Epstein EG, Sherman J, Blackman A, Sinkin RA. Testing the feasibility of Skype and FaceTime updates with parents in the neonatal intensive care unit. Am J Crit Care 2015 Jul;24(4):290-296 [FREE Full text] [doi: 10.4037/ajcc2015828] [Medline: 26134328]
31. Gulmans J, Vollenbroek-Hutten M, van Gemert-Pijnen L, van Harten W. A web-based communication system for integrated care in cerebral palsy: experienced contribution to parent-professional communication. Int J Integr Care 2012;12:e9 [FREE Full text] [Medline: 22977435]
32. Haney T, Tufts KA. A pilot study using electronic communication in home healthcare: implications on parental well-being and satisfaction caring for medically fragile children. Home Healthc Nurse 2012 Apr;30(4):216-224. [doi: 10.1097/NHH.0b013e31824c28f2] [Medline: 22456459]
33. Vismara LA, McCormick C, Young GS, Nadhan A, Monlux K. Preliminary findings of a telehealth approach to parent training in autism. J Autism Dev Disord 2013 Dec;43(12):2953-2969. [doi: 10.1007/s10803-013-1841-8] [Medline: 23677382]
34. Vismara LA, Young GS, Rogers SJ. Telehealth for expanding the reach of early autism training to parents. Autism Res Treat 2012;2012:121878 [FREE Full text] [doi: 10.1155/2012/121878] [Medline: 23227334]
35. 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]
36. Bradford NK, Armfield NR, Young J, Smith AC. Paediatric palliative care by video consultation at home: a cost minimisation analysis. BMC Health Serv Res 2014 Jul 28;14:328 [FREE Full text] [doi: 10.1186/1472-6963-14-328] [Medline: 25069399]
37. Braverman J, Samsonov DV. A study of online consultations for paediatric renal patients in Russia. J Telemed Telecare 2011;17(2):99-104. [doi: 10.1258/jtt.2010.100410] [Medline: 21163814]
38. Byczkowski TL, Munafo JK, Britto MT. Family perceptions of the usability and value of chronic disease web-based patient portals. Health Informatics J 2014 Jun;20(2):151-162. [doi: 10.1177/1460458213489054] [Medline: 24056751]
39. de Graaf M, Totte J, Breugem C, van Os-Medendorp H, Pasmans S. Evaluation of the compliance, acceptance, and usability of a web-based eHealth intervention for parents of children with infantile hemangiomas: usability study. JMIR Res Protoc 2013 Dec 17;2(2):e54 [FREE Full text] [doi: 10.2196/resprot.2897] [Medline: 24345450]

40. Hanlon-Dearman A, Edwards C, Schwab D, Millar MC, Longstaffe S. “Giving voice”: evaluation of an integrated telehealth community care model by parents/guardians of children diagnosed with fetal alcohol spectrum disorder in Manitoba. Telemed J E Health 2014 May;20(5):478-484. [doi: 10.1089/tmj.2013.0161] [Medline: 24568178]
41. Hopper B, Buckman M, Edwards M. Evaluation of satisfaction of parents with the use of videoconferencing for a pediatric genetic consultation. Twin Res Hum Genet 2011 Aug;14(4):343-346. [doi: 10.1375/twin.14.4.343] [Medline: 21787118]
42. Lee J, Du YE, Coki O, Flynn JT, Starren J, Chiang MF. Parental perceptions toward digital imaging and telemedicine for retinopathy of prematurity management. Graefes Arch Clin Exp Ophthalmol 2010 Jan;248(1):141-147. [doi: 10.1007/s00417-009-1191-6] [Medline: 19774354]
43. Mulgrew KW, Shaikh U, Nettiksimmons J. Comparison of parent satisfaction with care for childhood obesity delivered face-to-face and by telemedicine. Telemed J E Health 2011 Jun;17(5):383-387. [doi: 10.1089/tmj.2010.0153] [Medline: 21492028]
44. Nordfeldt S, Hanberger L, Berterö C. Patient and parent views on a Web 2.0 diabetes portal--the management tool, the generator, and the gatekeeper: qualitative study. J Med Internet Res 2010;12(2):e17 [FREE Full text] [doi: 10.2196/jmir.1267] [Medline: 20511179]
45. Wade SL, Walz NC, Carey JC, Williams KM. Brief report: description of feasibility and satisfaction findings from an innovative online family problem-solving intervention for adolescents following traumatic brain injury. J Pediatr Psychol 2009 Jun;34(5):517-522 [FREE Full text] [doi: 10.1093/jpepsy/jsn081] [Medline: 18667477]
46. Barlott T, Adams K, Díaz FR, Molina MM. Using SMS as a tool to reduce exclusions experienced by caregivers of people with disabilities in a resource-limited Colombian community. Disabil Rehabil Assist Technol 2015 Jul;10(4):347-354. [doi: 10.3109/17483107.2014.974223] [Medline: 25425412]
47. 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]
48. Royal College of Nursing. Adolescent transition care: RCN guidance for nursing staff 2013. 2016 Dec 01. URL: https://www2.rcn.org.uk/__data/assets/pdf_file/0011/78617/004510.pdf [accessed 2017-05-09] [WebCite Cache ID 6qKVRxj4m]
49. CareQualityCommission. From the pond into the sea: Children's transition to adult health services. 2014. URL: https://www.cqc.org.uk/sites/default/files/CQC_Transition%20Report.pdf [accessed 2017-12-21] [WebCite Cache ID 6vsG5HKqU]
50. National Institute for Health and Care Excellence. Transition from children's to adults' services- draft guideline. 2015. URL: https://www.nice.org.uk/guidance/ng43/documents/transition-from-childrens-to-adults-services-full-guideline2 [accessed 2017-12-21] [WebCite Cache ID 6vsG5cm5f]
51. Ganann R, Ciliska D, Thomas H. Expediting systematic reviews: methods and implications of rapid reviews. Implement Sci 2010 Jul 19;5:56 [FREE Full text] [doi: 10.1186/1748-5908-5-56] [Medline: 20642853]
52. Grover M, Williams C, Eisler I, Fairbairn P, McCloskey C, Smith G, et al. An off-line pilot evaluation of a web-based systemic cognitive-behavioral intervention for carers of people with anorexia nervosa. Int J Eat Disord 2011 Dec;44(8):708-715. [doi: 10.1002/eat.20871] [Medline: 22072408]
53. Sawyer SM, Afifi RA, Bearinger LH, Blakemore S, Dick B, Ezeh AC, et al. Adolescence: a foundation for future health. Lancet 2012 Apr 28;379(9826):1630-1640. [doi: 10.1016/S0140-6736(12)60072-5] [Medline: 22538178]
54. Ambresin A, Bennett K, Patton GC, Sanci LA, Sawyer SM. Assessment of youth-friendly health care: a systematic review of indicators drawn from young people's perspectives. J Adolesc Health 2013 Jun;52(6):670-681. [doi: 10.1016/j.jadohealth.2012.12.014] [Medline: 23701887]

Abbreviations

pre/post studies: pre- and postintervention uncontrolled studies
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
SMS: short message service

Edited by G Eysenbach; submitted 22.05.17; peer-reviewed by R Nightingale, M Reblin; comments to author 29.07.17; revised version received 15.09.17; accepted 04.11.17; published 05.01.18.

Please cite as:
Armoiry X, Sturt J, Phelps EE, Walker CL, Court R, Taggart F, Sutcliffe P, Griffiths F, Atherton H
Digital Clinical Communication for Families and Caregivers of Children or Young People With Short- or Long-Term Conditions: Rapid Review
J Med Internet Res 2018;20(1):e5
URL: http://www.jmir.org/2018/1/e5/
doi:10.2196/jmir.7999
PMID:29305339

Original Paper

Technological Health Intervention in Population Aging to Assist People to Work Smarter not Harder: Qualitative Study

Sonia Chien-I Chen[1,2], BA, MSc, PhD

[1]Connected Health Innovation Centre, Department of Leadership and Management, Ulster University, Newtownabbey, United Kingdom
[2]Ministry of Science and Technology, Taipei, Taiwan

Corresponding Author:
Sonia Chien-I Chen, BA, MSc, PhD
Connected Health Innovation Centre
Department of Leadership and Management
Ulster University
Shore Road
Newtownabbey, BT37 0QB
United Kingdom
Phone: 44 918192281
Email: drsoniachen@mail.com

Abstract

Background: Technology-based health care has been promoted as an effective tool to enable clinicians to work smarter. However, some health stakeholders believe technology will compel users to work harder by creating extra work.

Objective: The objective of this study was to investigate how and why electronic health (eHealth) has been applied in Taiwan and to suggest implications that may inspire other countries facing similar challenges.

Methods: A qualitative methodology was adopted to obtain insightful inputs from deeper probing. Taiwan was selected as a typical case study, given its aging population, advanced technology, and comprehensive health care system. This study investigated 38 stakeholders in the health care ecosystem through in-depth interviews and focus groups, which provides an open, flexible, and enlightening way to study complex, dynamic, and interactive situations through informal conversation or a more structured, directed discussion.

Results: First, respondents indicated that the use of technology can enable seamless patient care and clinical benefits such as flexibility in time management. Second, the results suggested that a leader's vision, authority, and management skills might influence success in health care innovation. Finally, the results implied that both internal and external organizational governance are highly relevant for implementing technology-based innovation in health care.

Conclusions: This study provided Taiwanese perspectives on how to intelligently use technology to benefit health care and debated the perception that technology prevents human interaction between clinicians and patients.

(J Med Internet Res 2018;20(1):e3) doi:10.2196/jmir.8977

KEYWORDS

health care; innovation; eHealth; technology; smart health; Taiwanese health care

Introduction

Rapid aging makes health care studies relevant and prevalent. This study is unique as it offers qualitative perspectives on how health providers can work smarter with technology, which differs from earlier studies on technology's ability to improve productivity through quantitative methods. This study extends the current knowledge of health care innovation by debating the perception of technology-led health care from the Taiwanese perspective.

An aging society implies a gradual decrease in a nation's productive labor force and an increase in its dependent population. Low fertility and mortality have led to a rapidly aging population in Taiwan. Consequently, the aging population has prompted people to work harder than ever. It is widely believed that diminishing caregiver support resources for increasing elderly populations, as witnessed in most developed countries, render many current care provision models for the elderly unsustainable [1-3].

Technology-driven systems have been proposed and promoted as potential solutions to help people to work *smarter* rather than harder. Despite the application of advanced technologies, concerns surrounding information governance and Taiwan's conservative attitude toward innovation and an actual loss of autonomy might hinder prosperous health care innovation [4]. This study aims to investigate how and why electronic health (eHealth) has developed in Taiwan and health care stakeholders' attitudes relative to innovation in a Taiwanese context to inspire other countries encountering similar issues.

Methods

Overview of Methods Used

To meet the study's aims and objectives, a qualitative methodology was adopted to obtain insightful inputs and facilitate deeper probing. Taiwan was selected as a typical case study, given its aging population, advanced technology, and comprehensive health care system. A qualitative methodology offers potential benefits that may inspire countries facing similar challenges by ensuring a better understanding of "how" and "why" eHealth has been implemented in Taiwan. This study investigated 38 stakeholders in the health care ecosystem through in-depth interviews and focus groups, which provides an open, flexible, and enlightening way to study complex, dynamic, and interactive situations through informal conversation or a more structured, directed discussion [5].

Data Collection

The recruitment of participants was voluntary through advertisement or word-of-mouth. The participants interviewed had given verbal consent according to ethical guidelines. This study was approved by the Ulster University's institutional review board. The participants included health care professionals, industry players, academic researchers, and government agents, and the sample was selected based on the literature review and *snow-ball* sampling. An open, flexible, and enlightening methodology facilitated the study of various complex, dynamic, and interactive situations. Interviews with participants were recorded and transcribed. The data collection, writing, and analysis progressed continuously in line with an interpretive research tradition [6]. In contrast to normal interviews, which resemble normal, direct conversations between an interviewer and interviewee [6,7], intensive interviews allow the interviewer to explore and examine a certain topic or an experience in detail; therefore, they are an effective method for interpretive inquiry. The interviewer aims to comprehend the topic while the interviewee owns the explored, pertinent experiences [8,9].

Thus, semistructured interviews that lasted up to 90 min were used to collect data. The questions included categories involving innovation and business management systems (BMS) in health care to understand the relationship between health care practice and business management. Additionally, the researchers observed the various personnel and organizational factors affecting the use of remote medical informatics to comprehend health care stakeholders' attitudes to technology acceptance and adaptability. The semistructured interview guide comprised the following several comprehensive exploratory aspects to identify how stakeholders' backgrounds, attitudes, competencies, and genders may relate to research outcomes:

- background information
- technology acceptance and adaptability
- levels of competency
- gender

The data transcription process followed the required reliability and validity procedures for qualitative studies.

Data Analysis

Data were collected and analyzed until theoretical saturation was reached. The study employed a thematic content analysis [10] of a qualitative analysis with recurring themes. A thematic analysis indicated the overall consistent themes that offered an in-depth understanding of these factors. Bryman [10] states that a qualitative content analysis is "probably the most prevalent approach to the qualitative analysis of documents" and "comprises a searching-out of underlying themes in the materials being analyzed." This was organized through deep familiarization with the data collected to further develop its major themes. These themes were then further analyzed to identify subthemes and structured to provide a comprehensive account.

According to Ruben and Babbie [11], content analysis is "essentially a coding operation," with coding referring to "the process of transforming raw data into a standardized form." According to Ryan and Bernard [12], a traditional content analysis "comprises techniques for reducing texts to a unit-by-variable matrix and analyzing that matrix quantitatively to test hypotheses;" the researcher can produce a matrix by applying a set of codes to a set of qualitative data assuming that the codes of interest were discovered and described beforehand.

A classical content analysis is essentially a quantitative method with a system of categories at its core and as a central tool. Consequently, the simplest type of evaluation involves counting the number of occurrences per category, assuming a relationship exists between the frequency of content and meaning. Moreover, more complex procedures can be used for the analysis of different indices that correlate two separate measurements and contingencies [13].

Theoretical Base

This study expands the current knowledge as to how technology is applied in health care innovation using a theoretical base including both eHealth [14-16] and innovation to explore ways to help people work smarter rather than harder [17-19]. Taiwan was selected as a case study because the country has exhibited significant eHealth performance and characteristics [1-3]. This study aims to explore the relationship between technology and working smart in health care to identify meaningful research gaps.

Definition of Work Smart

This research defines the term "work smart" as operating at a high level of efficiency and effectiveness. In contrast, the term "work hard" is a distinct and traditional labor-intensive work method that produces limited output. Technology might be smarter than manual effort, and these technological practices have purportedly replaced what has been regarded as “hard work.” The question is whether technology can help health care providers work smarter from different stakeholders’ perspectives.

Development of Technology Application in Health Care

Although technologies have been applied in health care for decades, the outcomes remain limited. Additionally, “eHealth” refers to health services facilitated by the Internet or related technologies at the intersection of medical informatics, public health, and business [15,16]. The topic of eHealth promise and performance has been broadly discussed since the year 2000 to promote effective health care [3,20-24].

Neuhauser and Kreps [3] investigated how eHealth can improve behavioral outcomes from an eHealth application perspective and noted that the outcomes vary according to diversity in human behavior. Pagliari [24] highlighted that a gap exists between user involvement and significant impacts from eHealth adoption by investigating eHealth design and evaluation and suggested that interdisciplinary collaboration between all potential stakeholders could produce more promising outcomes. According to Kreps and Neuhauser [20], eHealth processes hold significant potential concerning the accessibility of health information for both consumers and providers; consequently, this could enhance care quality by decreasing errors, increasing collaboration, and promoting health education.

Additionally, eHealth can be applied in communications to enhance the features of interactivity, multimodality, and mass customization as well as the opportunity for users to act as producers. Although challenges remain, eHealth is worthy of research in terms of its best use to improve accessibility for vulnerable populations as well as its long-term effects on public health [25]. However, Black et al [21] argued that eHealth has more benefits in theory than in practice, and its risk management must be considered to maximize its advantages.

Therefore, more empirical applications have been adopted with the development of eHealth that encompass health information technology, eHealth record systems, and health information exchange to produce comprehensive outcomes and patient engagement [23,24,26-29]. With an aging population, more research has focused on how technology, such as telemedicine, telehealth, and telecare, can be employed to remotely control chronic conditions.

Taiwan was selected as a typical case to study eHealth implementation because of its advancements in medicine and technology [30-35], its aging population phenomenon, and health care workforce shortage. Moreover, geographic gaps between Taiwan’s rural and urban populations indicate the need for research. Therefore, the Taiwanese Government sponsored several pilot schemes to manage issues related to population aging to help health care stakeholders work smarter.

Health care stakeholders are the individuals or groups directly or indirectly affected by eHealth systems; they either suffer from the problem that the eHealth system addresses or are affected by the eHealth solution itself. These stakeholders include eHealth users, health care professionals, patients, their family members, community social workers, and researchers.

Although telecare pilot schemes in Taiwan have exhibited positive results, current literature in most western countries suggests that the application of eHealth does not always translate into working smarter [36-38]. Additionally, literature provides evidence for professional resistance to the spread of innovation in health care [39-41]. Therefore, some stakeholders would prefer to remain in the safer “hard work” arena because of the uncertainty of innovation.

A review of the literature reveals that barriers to adoption are central to the lack of interoperability standards and care provider resistance. These barriers are based on ethical and security issues, including data protection, security, primary, and confidentiality issues [29,42-45]. This study reviews literature pertaining to disruptive technology to explore the status of Taiwan’s technology adoption in health care [17].

Results

Key Results

Figure 1 illustrates the relationships and collaborations among connected health stakeholders in the business ecosystem according to the data analysis results. Interviewees were selected and grouped into 8 categories based on this business ecosystem. The 38 interviews focused on following 5 main industries: software, hardware and electronics, Internet and telecommunications, health care, and total solutions. Current regulations in Taiwan have limited the practice of eHealth for hospitals and care institutions. Although it is difficult to define such institutions as academic or industry-related, they play an important role in this field. This research collects different perspectives with the aim of inspiring more creative approaches and exploiting the advantages of health care innovation.

The interviewees were labeled according to the 8 categories shown in Figure 1. One interviewee could represent more than one category because the company might be involved in various areas in the connected health ecosystem. Therefore, 34 subcategories were labeled from number 101 to number 805, as Table 1 indicates.

Data were analyzed based on the emergent patterns to respond to the research question. As there is diversity in the qualitative analysis because of the nature of richness and complexity in the qualitative research, the purpose of the analysis was to obtain rich information from respondents. A quasistatistical analysis was conducted using N-Vivo software (QSR International) to identify key themes with the adoption of a grounded approach in this analysis. In this study, codes were derived from the data through several steps, including data cleaning, data summarizing, data analysis, and data mining. First, data were screened through the integration process to merge different terms that have the same meaning, for example, connected health could be called remote health, telehealth, and telecare in the interviews; terms were merged according to the actual meaning of the interviewees. Second, data were summarized, clustered, and categorized based on the meaning of the interviewees.

For example, some interviewees mentioned the issue of Internet connection and wireless connection, which can be summarized and categorized into technology issues or infrastructure issues depending on their context. Third, data were analyzed and extracted according to the insightful meaning of interviewees as it is the stage of the data mining process.

The interview transcriptions were coded to generate summaries for translation into key themes through N-Vivo and the author's hand coding, as noted in Table 2. Summaries were further categorized into patient care, clinical benefits, and staff development. Additionally, leaders' visions, personalities, authority, and management capability were extracted as leadership and management themes. Finally, the hospital or internal institution's support system and the government or external institution's support system or ecosystem were classified into organizational issues.

Key Findings

The research outcomes can be classified as technology rationale, leadership management, and organizational governance, as shown in Table 2. First, respondents indicated that the use of technology could enable seamless patient care and provide clinical benefits including time management flexibility. Consequently, this bridges the gap between urban and rural access to the National Health Service. A technological learning curve could be overcome by staff development and the integration of younger and older generations. Other benefits, such as preventive medicine, have displayed decreased emergency room visit rates and provided both physical and mental total solutions to patients. One respondent stated the following:

> *The concept of preventive medicine through regular health recording is one of the crucial characteristics of remote health care.* [101, Senior Engineer]

Second, the results indicate that management's leadership and strategic thinking are the drivers of successful health innovation. One respondent stated the following:

> *Although remote areas in Taiwan have limited resources compared to the urban areas, they can provide the best model for the research and development of new [telehealth] products to enter the global market.* [403, Director]

Finally, this investigation implies that both internal and external organizational governance are highly relevant to the implementation of technology-based innovation in health care. Government funding is vital to encourage and initiate pilot schemes to develop comprehensive business models for business sustainability.

Figure 1. eHealth or connected health stakeholder relationship diagram.

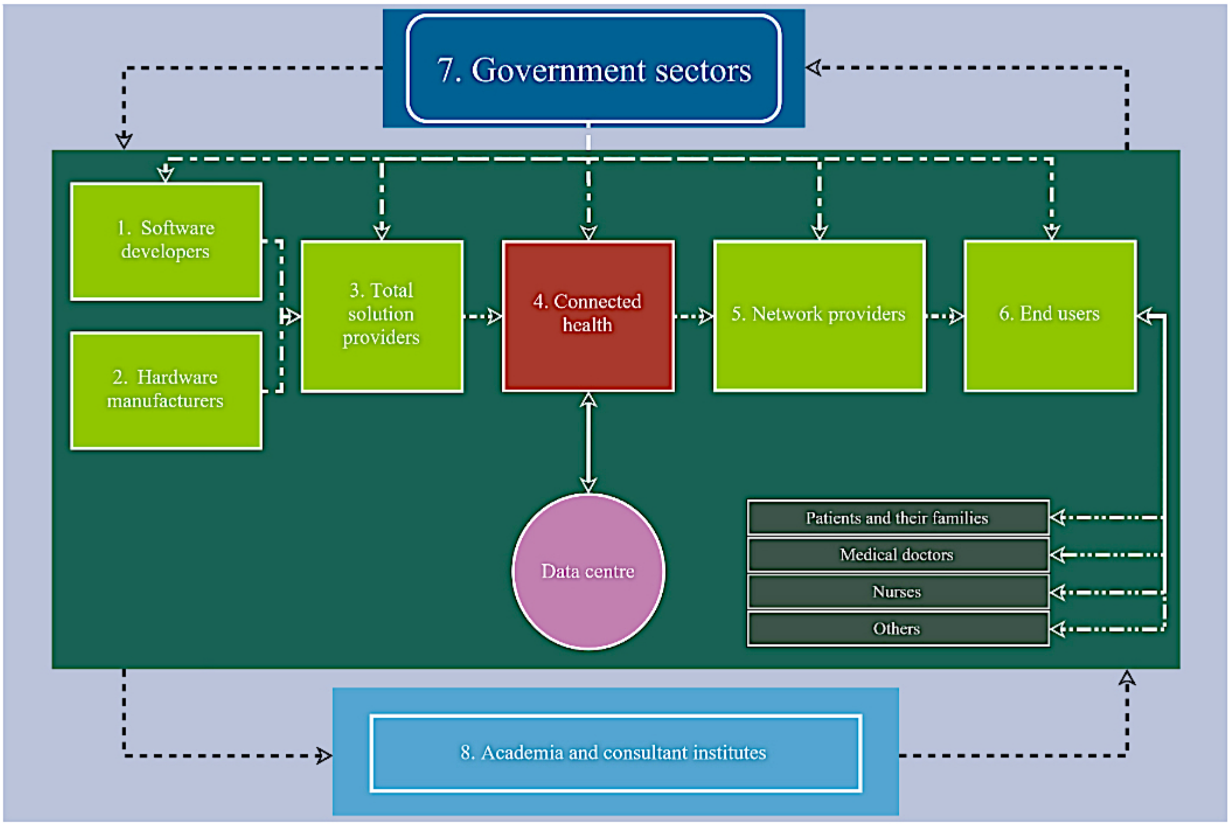


Table 1. Interviewees' categories, as summarized by the author.

Category and label number	Name
Software developers	
101	Far EasTone Telecommunications
102	Guidercare
Hardware manufacturers	
201	Netown Corporation
202	Far EasTone Telecommunications
203	Huede
204	Guidercare
205	Acomotech
Total solutions providers	
301	Netown Corporation
302	Far EasTone Telecommunications
303	Huede
304	Guidercare
305	Acomotech
Connected health care service providers	
401	Luo Dong Care Institute
402	En Chu Kong Hospital
403	Taoyuan Fu Hsing Township Health Station
404	Taipei Medical University Telehealth and Telecare Center
405	Mennonite Christian Hospital
406	Taiwan University Hospital Telehealth Center
407	Taiwan University Hospital
Network providers	
501	Far EasTone Telecommunications
End users	
601	National Taiwan University Hospital
602	Mennonite Christian Hospital
603	Taipei Medical University TH and Telecare Center
604	En Chu Kong Hospital
605	Luo Dong Care Institute
606	An-Tai Hospital
Government sectors	
701	Health and Welfare Department
702	Sang Chung Health Center
703	Luo Dong Care Institute
704	Taoyuan Fu Hsing Township Health Station
Academia	
801	National Taiwan University's Department of Engineering
802	Taipei University of Technology
803	Taipei University
804	Unified Life and Health

Category and label number	Name
805	National Taiwan University’s Department of Medicine

Table 2. Thematic output of the summarized data on eHealth/medical informatics, as summarized by the author.

Category	Summary from coding	Key themes
Patient care option	Seamless service	Technology rationale
	Infinite service	
	Connecting people	
	Preventive care	
	Personal information privacy	
	Public health education	
	Psychological and physical care	
	Quality	
	Efficacy	
Clinical benefits	Support for personnel shortages	
	Systematic and integrated patient records	
Staff development option	Cost-effective (improved medicine consumption and decreased medical costs)	
	Computer competency	
Leader’s vision	Short-sighted or insightful	Leadership and management
	Intellectual assets	
Leader’s personality	Work experience heritage	
Leader’s power or authority	Staff turnover	
	Creative capability	
	Innovation acceptance	
	Authority	
Leader’s management	Capability of execution	
	Risk management	
	Implementation management	
	Total solutions/package solution	
Government Hospital or ecosystem National health insurance	Financial	Organizational rationale
	Nonfinancial	

Discussion

Principal Findings

The key findings summarize the key themes into technology rationale, leadership management, and organizational governance, which emphasizes that innovation management in health care is as important as the advancement of technology itself. To sustain health care businesses, the tight collaboration among stakeholders in government, industry, and academia is significant. This section discusses the core rationales influencing health care stakeholders in adapting technology to their objectives in the interview discourse and the literature.

Technological Rationale

Concerns with personal information security must be addressed before designing eHealth care systems; however, if the benefits can outweigh the concerns, these systems may become as popular as online banking in the future. One respondent stated the following:

> *I believe that someday, accessing health information will be like online banking and online shopping.* [406, Head Nurse]

The securing of personal medical data should not be perceived as an insurmountable problem that prevents a telecare system’s adoption and development. Technology has the aspiration to create a mutually beneficial situation that propels the following ideas: patients will obtain better health outcomes, and physicians will experience more achievements and satisfaction.

Electronic data can be easily integrated with existing hospital information systems and, ultimately, contribute by decreasing financial burdens. More systematic and integrated patient records can be developed to offer better health services and decrease diagnostic errors, ensuring quality and cost-effective benefits. One respondent stated the following:

> *The blood sugar levels of those who have joined the telecare project have become more controlled* [402, Senior Nurse]

Mechanisms to develop innovative information technology–enabled services are particularly important [46]. This research posits that the use of technologies to improve patient-centered data management has increasingly become a research area of focus [47,48].

Additionally, this research noted one significant theme: respondents described technology in health care as a crucial tool for *connecting* people rather than *replacing* them. Technology can be used to bridge gaps, specifically, the isolation between older and younger generations and also rural and urban societies. One respondent stated the following:

> *Technology plays a role in delivering humanity and care to people; a remote environment causes distance among people, but technology connects them.* [602, MIS Director]

Furthermore, clinicians were concerned with the accuracy of biometric data. The respondents addressed these concerns because they claimed that they could typically collect more accurate data than hospitals by eliminating “white coat syndrome.”

Most respondents endorsed the necessity of tele-health care systems, yet the question of how to manage them seems to be an issue in need of exploration. One respondent stated the following:

> *Telehealth is necessary, and how you use it will really make a difference.* [403, Director]

The cost of adapting technology might initially pose problems because of growth, but upgrading and renewal costs become more cost-effective in the long term. A greater acceptance of innovation is possible by objectively demonstrating its benefits and shortcomings to the public. Ultimately, an eHealth system’s benefits will outweigh any potential problems in the long term.

Leadership Management

The respondents noted that a leader’s vision, personality, authority, and management style are highly relevant to a tele-health care project’s success.

Employees might take their innovative ideas to rival companies if their talents are not appreciated. Many cases have proven that leadership plays an important role, and respondents note that poor management causes high staff turnover and the loss of companies’ intelligent, intangible assets. One respondent stated the following:

> *Our previous manager did not care about innovation, but short-term profits [...] Now, our competitors have taken opportunities and have already scored some achievements by recruiting our old staff with innovative ideas…* [101, Senior Engineer]

Staff development activities could help some staff members absorb global trends, which might then encourage and motivate them to accept innovation. This might improve their adaptability to new technology. One respondent stated the following:

> *I will not consider the learning of new technology as an issue, but an opportunity for staff blending.* [405, Head Nurse]

Leadership attitude is a driver of eHealth success in Taiwan with collaborations among industry, government, and academia. Furthermore, literature suggests that health care lacks creativity; however, the interview findings support the assertion that leaders’ creative thinking and capability for innovation are beneficial to innovation success [49].

Many people have low innovation acceptance because of limited capability to adapt. An effective leader should be cognizant of this and properly manage the situation to further develop comprehensive staff communication accordingly. One respondent stated the following:

> *Any new system will have an unavoidable run-in period…You should inform people about the potential risks during the operation of a pilot scheme and have a backup plan.* [403, Director]

Physician support is significant because physicians have authority over patients as well as their trust. Furthermore, physician recommendations are more convincing to patients than recommendations from other sources. One respondent stated the following:

> *Physicians are more convincing in promoting tele-health care systems, as patients perceive them as having authority.* [406, Director]

Organizational Rationale

Although an outstanding leader is important, the successful implementation of technology in health care relies on both individuals and organizations. Government funding and infrastructure expenses along with organizational support will be beneficial for both clinicians and companies when transforming innovative ideas into concrete practices. One respondent stated the following:

> *The government must take the lead for the first step by offering funding; otherwise, we cannot go anywhere…* [101, Senior Engineer]

Bureaucracy and inflexibility can be a disadvantage in these government-funded schemes. One respondent stated the following:

> *The public sector has this type of issue, as all projects must be reclaimed within an annual calendar period [...] but some results cannot be shown within one to two years...* [403, Director]

Regarding financial sustainability in health care, collaborating with unlimited needs but limited financial capability has become the norm for governments, and political calculations are often labeled as unrealistic. However, health care provisions impact everyone's lives, and any cost-cutting plans executed by the government or employers will have a significant effect on people [50,51].

The method of integrating a resource and its ecosystem is relevant to how organizations can benefit from leadership management. Some organizations have used their knowledge, skills, and experience in exchange for assets they did not have, such as devices or networks. Moreover, the government typically has the authority to coordinate different departments and organizations. Nevertheless, a business model that considers what occurs after government sponsorship should be developed.

Conclusions

This study dealt with the concept of "work smarter," not work harder, by examining technology-driven systems from a qualitative perspective. It contributed to current knowledge by debating the perception that technology in health care prevents human interaction between clinicians and patients. The originality of this study lies in the focus on the relationship between technology and its relevant stakeholders in health care, which differs from the focus of recent literature that emphasizes the productivity and performance of technology. Additionally, this study extends the knowledge from technology-focused innovation to health business innovation through a qualitative lens. The Taiwanese health care system offers compelling information to support the claim that technology can be a tool to connect all stakeholders in the health care ecosystem.

The conclusion summarizes three themes for practical implications. First, the technology user should control technology through designing that serves rather than hinders. Second, leaders should manage innovation through creativity for resource integration and acquisition, and its long-term benefits should be emphasized over its short-term advantages. Third, although government funding and organizational support are relevant to health care innovation, appropriate BMS should be developed for long-term success after the government ceases its support. Interview respondents confirm that technology is beneficial when encouraging health providers to work smarter if it is used to reinforce organizations' resources, leadership management, and collaboration with their ecosystems. To achieve this goal, the findings of this study strongly recommend health care stakeholders consider creating competitive BMS to retain business sustainability.

Although the qualitative research in a regional study can be a limitation to the research validity, it can be considered and developed in a future study. The findings emphasize how users should employ technology to enable innovation rather than constrain innovation while pursuing its novelty. Technology users should work "smarter" when using technology to add value to its adaption. Otherwise, technology will compel users to work harder by creating additional work.

Acknowledgments

This research was partially supported by the Vice Chancellor Research Studentship from Ulster University. The author would like to thank her supervisors, Professor Rodney McAdam, Dr Michael Brennan, and Dr Shirley Davey from Ulster University for sharing their pearls of wisdom during this research. The author is also immensely grateful to Mr Francis Kane, Mr Ray Palmer, and Dr Radwan Kharabsheh for their comments on an earlier version of the manuscript.

Conflicts of Interest

None declared.

References

1. Liu W, Huang C, Wang C, Lee K, Lin S, Kuo H. A mobile telephone-based interactive self-care system improves asthma control. Eur Respir J 2011 Feb;37(2):310-317 [FREE Full text] [doi: 10.1183/09031936.00000810] [Medline: 20562122]
2. Blaya JA, Fraser HS, Holt B. E-health technologies show promise in developing countries. Health Aff (Millwood) 2010 Feb;29(2):244-251 [FREE Full text] [doi: 10.1377/hlthaff.2009.0894] [Medline: 20348068]
3. Neuhauser L, Kreps GL. Rethinking communication in the e-health era. J Health Psychol 2003 Jan;8(1):7-23. [doi: 10.1177/1359105303008001426] [Medline: 22113897]
4. Stowe S, Harding S. Telecare, telehealth and telemedicine. Eur Geriatr Med 2010 Jun 30;1(3):193-197. [doi: 10.1016/j.eurger.2010.04.002]
5. Walsham G. Doing interpretive research. Eur J Inf Syst 2006;15(3):320-330. [doi: 10.1057/palgrave.ejis.3000589]
6. Lofland J, Lofland L. Analyzing Social settings. A Guide to Qualitative Observation and Analysis. Belmont, CA: Wadsworth; 1984.
7. Fontana A, Frey J. The art of science. In: Handbook of Qualitative Research. Thousand Oaks, CA: Sage Publications; 1994:361-376.
8. Marshall C, Rossman G. Designing Qualitative Research. Thousand Oaks, CA: Sage Publications; 2014.
9. Kharabsheh R. Critical success factors of technology parks in Australia. Int J Econ Finance 2014 Dec 16;4(7):57. [doi: 10.5539/ijef.v4n7p57]
10. Bryman A. The Disneyization of Society. Thousand Oaks, CA: Sage Publications; 2004.

11. Rubin A, Babbie E. Empowerment series: Research Methods for Social Work. Pacific Grove, CA: Brooks Cole; 2016.
12. Ryan GW, Bernard HR. Data management and analysis methods. In: Denzin NK, Lincoln YS, editors. The SAGE Handbook of Qualitative Research. Florida: Sage Publications, Inc; 2000:769-802.
13. Titscher S. Methods of Text and Discourse Analysis: In Search of Meaning. London: Sage Publications; 2000.
14. Eysenbach G. What is e-health? J Med Internet Res 2001;3(2):E20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
15. 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 09;324(7337):573-577 [FREE Full text] [Medline: 11884321]
16. Christensen C, Grossman J, Hwang J. The Innovator's Prescription:A Disruptive Solution for Healthcare. New York: McGraw-Hill; 2009.
17. Nassiri N, Natarajan S, Margolis D, Marks L. Targeted prostate biopsy: lessons learned amidst the evolution of a disruptive technology. Urology 2015;86(3):432-438. [doi: 10.1016/j.urology.2015.07.001]
18. Lauer M, D'Agostino SR. The randomized registry trial--the next disruptive technology in clinical research? New England J Med 2013 Oct 24;369(17):1579. [doi: 10.1056/NEJMp1310102] [Medline: 23991657]
19. Lee M. From recipient to donor: How Taiwan transformed its healthcare system. Japan Med Assoc J 2011;55(1):23-25.
20. Kreps G, Neuhauser L. New directions in eHealth communication: opportunities and challenges. Pat Edu and Couns 2010;78(3):329-336. [doi: 10.1016/j.pec.2010.01.013] [Medline: 20202779]
21. Black A, 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;8(1):e1000387. [doi: 10.1371/journal.pmed.1000387] [Medline: 21267058]
22. Smith S, Koppel R. Healthcare information technology's relativity problems: a typology of how patients' physical reality, clinicians' mental models, and healthcare information technology differ. J Amer Med Informatics Assoc 2014;21(1):117-131. [doi: 10.1136/amiajnl-2012-001419] [Medline: 23800960]
23. Nundy S, Dick J, Chou C, Nocon R, Chin M, Peek M. Mobile phone diabetes project led to improved glycemic control and net savings for Chicago plan participants. Health Aff (Millwood) 2014;33(2):265-272. [doi: 10.1377/hlthaff.2013.0589] [Medline: 24493770]
24. Pagliari C. Design and evaluation in eHealth: challenges and implications for an interdisciplinary field. J Med Internet Res 2007 May 27;9(2):e15 [FREE Full text] [doi: 10.2196/jmir.9.2.e15] [Medline: 17537718]
25. Ahern D, Kreslake J, Phalen J. What is eHealth (6): Perspectives on the evolution of eHealth research. J Med Internet Res 2006;8(1):e4 [FREE Full text] [doi: 10.2196/jmir.8.1.e4]
26. Yasnoff W, Sweeney L, Shortliffe E. Putting health IT on the path to success. J Am Med Assoc 2013;309(10):989-990. [doi: 10.1001/jama.2013.1474] [Medline: 23483170]
27. Furukawa M, King J, Patel V, Hsiao C, Adler-Milstein J, Jha A. Despite substantial progress in EHR adoption, health information exchange and patient engagement remain low in office settings. Health Aff (Millwood) 2014 Aug 7;33(9):1672-1679. [doi: 10.1377/hlthaff.2014.0445] [Medline: 25104827]
28. Koch S, Vimarlund V. Critical advances in bridging personal health informatics and clinical informatics. Yearb Med Inform 2012;7(1):48-55. [Medline: 22890341]
29. Pare G, Jaana M, Sicotte C. Systematic review of home telemonitoring for chronic diseases: the evidence base. J Amer Med Informatics Assoc 2007;14(3):269-277. [doi: 10.1197/jamia.M2270]
30. Lin C. Telecare projects in Taiwan. Gerontechnology 2010;9(3):400-407. [doi: 10.4017/gt.2010.09.03.003.00]
31. Chang R, Lin S, Aron D. A pay-for-performance program in Taiwan improved care for some diabetes patients, but doctors may have excluded sicker ones. Health Aff (Millwood) 2012 Jan 1;31(1):93-102. [doi: 10.1377/hlthaff.2010.0402] [Medline: 22232099]
32. Chang N, Huang C, Yang Y. The effects of economic crises on health care. Hu Li Za Zhi 2010 Aug;57(4):89-94. [Medline: 20661861]
33. Coughlin J, Pope J. Innovations in health, wellness, and aging-in-place. IEEE Eng Med Biol Mag 2008;27(4):47-52. [doi: 10.1109/MEMB.2008.919493]
34. Beard J, Biggs S, Bloom D, Fried L, Hogan P, Kalache R, et al. Weforum.: World Economic Forum; 2012. Global Population Aging: Peril or promise? URL: https://www.weforum.org/reports/global-population-ageing-peril-or-promise [accessed 2017-12-18] [WebCite Cache ID 6vnSPEsKL]
35. Ricciardi L, Mostashari F, Murphy J, Daniel J, Siminerio E. A national action plan to support consumer engagement via e-health. Health Aff (Millwood) 2013 Feb 1;32(2):376-384. [doi: 10.1377/hlthaff.2012.1216] [Medline: 23381531]
36. Greenhalgh T, Procter R, Wherton J, Sugarhood P, Shaw S. The organising vision for telehealth and telecare: discourse analysis. BMJ Open 2012;2(4):e001574. [doi: 10.1136/bmjopen-2012-001574]
37. Proctor JJ, Proctor IJ. Google.tl. 2012 Jan 3. Visual identification information used as confirmation in a wireless communication URL: http://www.google.tl/patents/US8090616 [accessed 2017-12-18] [WebCite Cache ID 6vnShiq44]

38. Rahurkar S, Vest J, Menachemi N. Despite the spread of health information exchange, there is little evidence of its impact on cost, use, and quality of care. Health Aff (Millwood) 2015 Mar 1;34(3):477-483. [doi: 10.1377/hlthaff.2014.0729] [Medline: 25732499]
39. Perle J, Langsam L, Nierenberg B. Controversy clarified: an updated review of clinical psychology and tele-health. Clin Psychol Rev 2011 Dec 31;31(8):1247-1258. [doi: 10.1016/j.cpr.2011.08.003] [Medline: 21963670]
40. Dillon E, Loermans J. Telehealth in Western Australia: the challenge of evaluation. J Telemed Telecare 2003;9(Suppl 2):S15-S19. [doi: 10.1258/135763303322596147] [Medline: 14728750]
41. Weitzman E, Cole E, Kaci L, Mandl K. Social but safe? Quality and safety of diabetes-related online social networks. J Amer Med Informatics Assoc 2011;18(3):292-297. [doi: 10.1136/jamia.2010.009712] [Medline: 21262920]
42. Wynia M, Dunn K. Dreams and nightmares: practical and ethical issues for patients and physicians using personal health records. J Law Med Ethics 2010;38(1):64-73. [doi: 10.1111/j.1748-720X.2010.00467.x] [Medline: 20446985]
43. Cushman R, Froomkin AM, Cava A, Abril P, Goodman KW. Ethical, legal and social issues for personal health records and applications. J Biomed Inform 2010 Oct;43(5 Suppl):S51-S55 [FREE Full text] [doi: 10.1016/j.jbi.2010.05.003] [Medline: 20937485]
44. Peek S, Wouters E, van Hoof J, Luijkx K, Boeije H, Vrijhoef H. Factors influencing acceptance of technology for aging in place: a systematic review. Intl J Med Informatics 2014;83(4):235-248. [doi: 10.1016/j.ijmedinf.2014.01.004] [Medline: 24529817]
45. Yang H, Hsiao S. Mechanisms of developing innovative IT-enabled services: a case study of Taiwanese healthcare service. Technovation 2009;29(5):327-337. [doi: 10.1016/j.technovation.2009.01.006]
46. Nicolini D, Powell J, Conville P, Martinez-Solano L. Managing knowledge in the healthcare sector: a review. Intl J Mgmt Rev 2008;10(3):245-263. [doi: 10.1111/j.1468-2370.2007.00219.x]
47. Hall A, Walton G. Information overload within the health care system: a literature review. Health Info Libraries J 2004;21(2):102-108. [doi: 10.1111/j.1471-1842.2004.00506.x] [Medline: 15191601]
48. Herzlinger RE. Why innovation in health care is so hard. Harv Bus Rev 2006 May;84(5):58-66, 156. [Medline: 16649698]
49. Bonoli G. Political institutions, veto points, and the process of welfare state adaptation. In: The New Politics of the Welfare State. Oxford: Oxford University Press; 2001.
50. Rhodes M. The political economy of social pacts: Competitive corporatism and European welfare reform. In: The New Politics of the Welfare State. Oxford: Oxford University Press; 2001.
51. Barlow J, Singh D, Bayer S, Curry R. A systematic review of the benefits of home telecare for frail elderly people and those with long-term conditions. J Telemed Telecare 2007;13(4):172-179. [doi: 10.1258/135763307780908058] [Medline: 17565772]

Abbreviations

BMS: business management systems
eHealth: electronic health

Edited by G Eysenbach; submitted 14.09.17; peer-reviewed by V Srica, R Pankomera; comments to author 18.10.17; revised version received 30.10.17; accepted 30.10.17; published 04.01.18.

Please cite as:
Chen SCI
Technological Health Intervention in Population Aging to Assist People to Work Smarter not Harder: Qualitative Study
J Med Internet Res 2018;20(1):e3
URL: http://www.jmir.org/2018/1/e3/
doi:10.2196/jmir.8977
PMID:29301736

Original Paper

Unhappy Patients Are Not Alike: Content Analysis of the Negative Comments from China's Good Doctor Website

Wei Zhang[1], PhD; Zhaohua Deng[1], PhD; Ziying Hong[1], MA; Richard Evans[2], PhD; Jingdong Ma[1], PhD; Hui Zhang[3], PhD

[1]Institute of Smart Health, School of Medicine and Health Management, Huazhong University of Science and Technology, Wuhan, China
[2]Department of Business Information Management and Operations, University of Westminster, London, United Kingdom
[3]School of Public Administration, Guangzhou University, Guangzhou, China

Corresponding Author:
Zhaohua Deng, PhD
Institute of Smart Health
School of Medicine and Health Management
Huazhong University of Science and Technology
No. 13 Hangkong Road, Qiaokou District
Wuhan, 430030
China
Phone: 86 15926318828
Email: zh-deng@hust.edu.cn

Abstract

Background: With the rise in popularity of Web 2.0 technologies, the sharing of patient experiences about physicians on online forums and medical websites has become a common practice. However, negative comments posted by patients are considered to be more influential by other patients and physicians than those that are satisfactory.

Objective: The aim of this study was to analyze negative comments posted online about physicians and to identify possible solutions to improve patient satisfaction, as well as their relationship with physicians.

Methods: A Java-based program was developed to collect patient comments on the Good Doctor website, one of the most popular online health communities in China. A total of 3012 negative comments concerning 1029 physicians (mean 2.93 [SD 4.14]) from 5 highly ranked hospitals in Beijing were extracted for content analysis. An initial coding framework was constructed with 2 research assistants involved in the codification.

Results: Analysis, based on the collected 3012 negative comments, revealed that unhappy patients are not alike and that their complaints cover a wide range of issues experienced throughout the whole process of medical consultation. Among them, physicians in Obstetrics and Gynecology (606/3012, 20.12%; P=.001) and Internal Medicine (487/3012, 16.17%; P=.80) received the most negative comments. For negative comments per physician, Dermatology and Sexually Transmitted Diseases (mean 5.72, P<.001) and Andrology (mean 5, P=.02) ranked the highest. Complaints relating to insufficient medical consultation duration (577/3012, 19.16%), physician impatience (527/3012, 17.50%), and perceived poor therapeutic effect (370/3012, 12.28%) received the highest number of negative comments. Specific groups of people, such as those accompanying older patients or children, traveling patients, or very important person registrants, were shown to demonstrate little tolerance for poor medical service.

Conclusions: Analysis of online patient complaints provides an innovative approach to understand factors associated with patient dissatisfaction. The outcomes of this study could be of benefit to hospitals or physicians seeking to improve their delivery of patient-centered services. Patients are expected to be more understanding of overloaded physicians' workloads, which are impacted by China's stretched medical resources, as efforts are made to build more harmonious physician-patient relationships.

(J Med Internet Res 2018;20(1):e35) doi:10.2196/jmir.8223

KEYWORDS

patient satisfaction; physician-patient relationship; Good Doctors website; patient complaint.

Introduction

Benefit of Web-Based Patient Complaints

The identification and recording of patient complaints is vital for improving the quality of health care services and maintaining good physician-patient relationships [1]. However, not all patients complain when they are dissatisfied. Previous studies [2,3] on quality management have revealed that only one-third of patients complain when they experience unsatisfactory service. Reasons for patients not complaining include lack of contact information for customer complaints offices, complicated complaint procedures, and insufficient transparency in dealing with complaints [4-6]. With the growth in worldwide Internet availability and usage, a behavioral shift is identified where people are moving from traditional offline complaint channels to expressing their views in relation to unsatisfactory health care experiences via the Internet. Such channels provide an alternative approach for patients to discuss their substandard medical experiences and share them with others, especially when they feel the service provider failed to take effective action [7]. Compared with on-site complaints, patients benefit from increased time and description allowance when complaining online. Salma et al [8] conducted a qualitative study into patients' attitudes toward submitting Web-based feedback and ratings to general practitioners (GPs) in England; they suggested that patients leave comments online mainly for one of the following three reasons: (1) the ability and ease of giving it remotely, (2) availability to the public, and (3) the perceived serious attitude of the GPs toward the Web-based comments.

Web-Based Patient Comments and Their Impact

Undoubtedly, comments posted online by dissatisfied patients exert a potentially greater impact on reviewer behaviors relevant to medical scenarios. Most of the time, these comments could be identified in specific physician rating websites (PRWs) where patients can share their medical experiences with a certain physician, serving as a reference for medical decision making [9]. PRWs enable patients to post comments relating to their experiences and emotions following appointments online [10], which identifies patients' actual experiences of health care providers. With this type of information sharing, not only can physicians be better informed about patient concerns [11], but they can also benefit from being able to offer the most appropriate physicians and facilitating the communication between them. Previous research [12] has concluded that comments for physicians are positive in nature. A content analysis of 3000 narrative comments posted on a German PRW revealed that the majority of comments were positive (2400/3000, 80%), with 16% (480/3000) being negative and 4% (120/3000) neutral [13]. Similar results were found on the largest doctor review website in China, the Good Doctor [14]. Other researchers have further explored PRWs via new analytical techniques [15-20]. Greaves et al [16] adopted a machine-learning approach to classify 6421 comments obtained from the English National Health Service website, splitting them into positive and negative posts. Another study [18] revealed that complaints relating to access to appointments, appointment waiting times, and time spent with a physician were viewed as most important. Li et al [21] examined the proportion and position of Web-based negative reviews and their effect on patient decision making. They found an increase in the number of negative reviews and identified that the higher the position of a negative review, the greater was the reluctance of patients to use a physician's services. It is evident that patients favor physicians who receive the most positive comments over those with negative or dissatisfied comments. However, negative or dissatisfied comments can hardly be avoided, considering the unpredictable behavior of individual patients following their health care experience. Thus, readers of Web-based reviews must be cautious when encountering negative or dissatisfied comments when making their judgments.

Negative comments posted may harm or create a damaging image of physicians and increase the dissatisfaction of patients toward the physician or hospital [22]. A study of GPs in England suggests that Web-based negative comments may affect a GP's confidence and self-esteem and lead to self-defense during their future work [23]. Extant studies [9,10,12,14,18] have predominantly focused on large-scale investigations of PRWs and the overall attitudes of comment providers, with little attention being given to the detail of negative comments posted. This study aims to analyze factors associated with negative comments posted online and provide empirical evidence for the understanding of unhappy patients and their comments. The research questions posed in this investigation include the following:

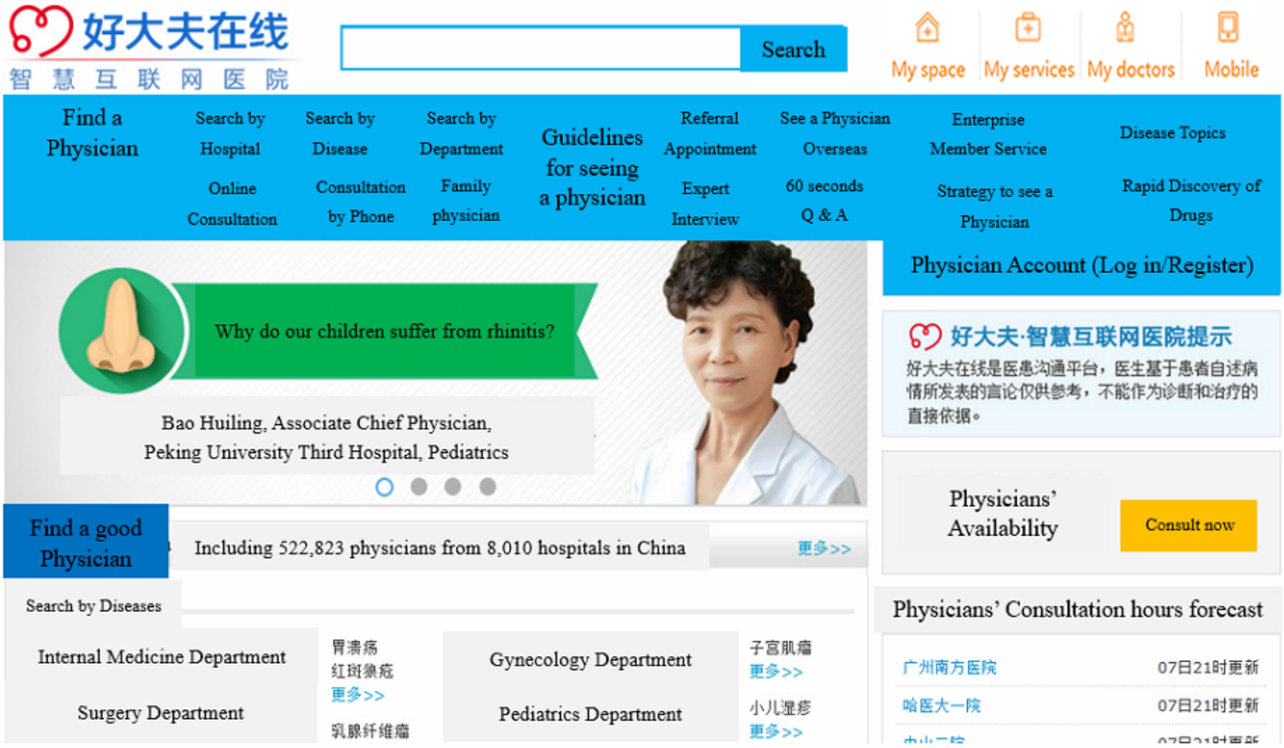
> RQ1: What do patients complain about online in relation to their physicians?
>
> RQ2: How can we improve the physician-patient relationship through the lens of Web-based negative comments?

Methods

Data Collection

Data were collected from the Good Doctor website, one of China's largest online physician-patient communication platforms (see Figure 1 for home page visual). The platform was established in 2006 and currently has 7794 hospitals and over 500,000 physicians registered to the service [24]. The website enables patients to access three types of services: (1) health information search, (2) medical consultation, and (3) patient feedback. For the health information search, users are able to obtain medical knowledge, medical news, and expert opinions. For the medical consultation service, patients can consult specific physicians via picture, telephone, or videoconference. In addition, online booking of appointments for referrals, delivery of drugs, arrangement of remote medical consultations, and the issuing of electronic prescriptions are also possible. For patient feedback, patients can rate their physicians and submit comments. As such, Good Doctor website is the first nationwide online platform for patients to share their experiences with physicians in China. Feedback mechanisms allow patients to vote, comment, write thank-you letters, and send electronic gifts to their physicians. The website's rating system automatically recommends good doctors based on patient ratings and provides a comprehensive evaluation of the expertise of physicians, as well as a hospital's reputation.

Figure 1. The home page of the Good Doctor website (accessed September 8, 2017).



好大夫·智慧互联网医院提示

好大夫在线是医患沟通平台，医生基于患者自述病情所发表的言论仅供参考，不能作为诊断和治疗的直接依据。

The Good Doctor website declares that all personal patient information contained on their website is removed automatically, including medical profiles and consultation records. All the content generated by the website's users and doctors is regarded as shared property between its users, doctors, and the website. In addition, content is made publicly available for the common good, but commercial use is prohibited.

Given that China's medical resources are unevenly distributed, economic prosperity and administrative powers always imply that abundant medical services are available. In this study, we focus solely on Beijing, China's capital, where the leading and preeminent hospitals in China are situated. As a pilot study, comments were collected on physicians from 5 reputable tertiary referral hospitals in Beijing, including China-Japan Friendship Hospital, Chinese PLA General Hospital, Peking University First hospital, Peking University Third Hospital, and Peking Union Medical College Hospital. A predesigned Java-based Web crawler was utilized to obtain all webpages concerning the 5 hospitals on Good Doctor website in September 2016. Data collected were stored in a MySQL database through a process of page parsing and information extraction. Data consisted of 140,591 patient comments relating to 5727 physicians posted on the physicians' home page.

On the Good Doctor website, patients can rate their satisfaction with medical services provided on the following scale: very satisfied, satisfied, ordinary, no comment, and dissatisfied; this rating relates to the physician's attitude and therapeutic effect in his or her latest consultation. As our research concerns patients' dissatisfaction, we have selected patient comments where at least one item (physician attitude or therapeutic effect) was considered to be unsatisfactory. After filtering "noisy" and null data by text preprocessing, we obtained 3012 dissatisfied comments (see Table 1 for details). Among them, 1565 (51.96%, 1565/3012) patients rated their experience as unsatisfactory in relation to the physician's attitude and manner. For the remaining comments, 67 out of 1447 (4.63%) indicated that they were either very satisfied or satisfied with the physician's attitude; in terms of therapeutic effect, the figure was 197 (12.59%, 197/1565), which is much higher than that recorded for "physician's attitude." We compared the two groups of negative comments toward physician attitude and manner, and no statistical difference was found (P=.90). All comments concerned 1029 physicians in total (mean 2.93 [SD 4.14]). The average length of a comment is 195.83 words (min=13, max=3345, [SD 189.19]). For each physician, the number of negative comments they received ranged from 1 to 69. After grouping patients using the classification of the Good Doctor website, we found that the Department of Obstetrics and Gynecology received the most negative comments with 606 (P=.001), followed by the Department of Internal Medicine with 487 (P=.08). Meanwhile, for the number of negative comments per physician, the Department of Dermatology and Sexually Transmitted Diseases (STD) and Andrology ranked the highest with 5.72 and 5 comments per physician, respectively. Table 2 provides further details.

Table 1. Overview of the data.

Dissatisfied items	Very satisfied	Satisfied	Ordinary	No comment	Dissatisfied	*F* value	Significance
Attitude	28	39	224	579	1565	0.02	.90
Effect	63	134	355	25	1565		

Table 2. Number of negative comments by departments.

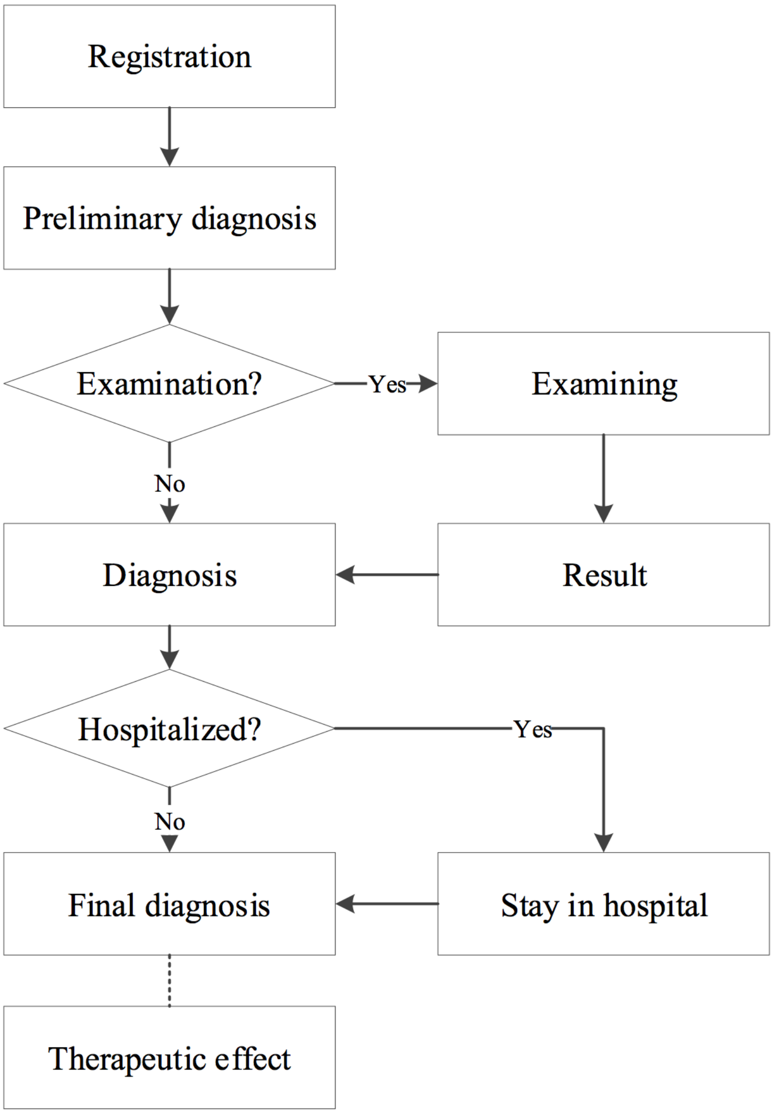
Department	Number of negative comments	Number of physicians	Negative comments per physician	Standard deviation	*P* value
Internal Medicine	487	204	2.39	3.11	.08
Surgical Department	356	164	2.17	1.99	.02
Obstetrics and Gynecology	606	132	4.59	7.24	.001
Reproductive Center	103	23	4.48	5.25	.08
Pediatrics	104	45	2.31	2.00	.32
Orthopedics Surgery	162	64	2.53	2.83	.45
Ophthalmology	68	42	1.62	0.99	.04
Stomatology	75	39	1.92	1.86	.13
Otorhinolaryngology (ear, nose, and throat) and Head and Neck	128	41	3.12	2.52	.77
Oncology	28	11	2.55	1.75	.76
Dermatology and Sexually Transmitted Diseases	372	65	5.72	6.76	<.001
Andrology	125	25	5.00	6.73	.02
Psychiatry	10	5	2.00	1.41	.62
Traditional Chinese Medicine	90	38	2.37	1.52	.41
Integrated Traditional Chinese Medicine and Western Medicine	4	4	1.00	0.00	.35
Intervention Therapy	8	4	2.00	2.00	.65
Rehabilitative Medicine	10	3	3.33	4.04	.87
Sports Medicine	31	19	1.63	0.76	.17
Anesthesiology	19	9	2.11	3.33	.56
Occupational Diseases	1	1	1	—	—
Medical Imaging	14	10	1.40	1.27	.24
Others	211	81	2.60	2.41	.48

Data Analysis

Content analysis was used to explore the factors associated with patient dissatisfaction [25,26]. Reader et al [27] developed a coding taxonomy for patient complaints through a systematic review of 59 studies. Following thematic analysis and grouping, they conceptualized three distinct domains of complaint: (1) safety and quality of the clinical care received, (2) the management of health care organizations, and (3) problems associated with health care staff-patient relationships. On the basis of their framework and the workflow of medical consultation in China (see Figure 2), we developed a coding framework with five dimensions of patient dissatisfaction that occurred before, during, or after the patient received health care from a specific physician, consisting of (1) the physician's attitude, (2) therapeutic effect, (3) ignorance of patient, (4) limited treatment time, and (5) misconduct or bad attitude of the nurse and/or other staff.

Patients and/or their caregivers submit comments online to advise those seeking health care from doctors when experiencing similar diseases or symptoms. Generally, negative comments can be divided into three types: (1) content, including an explanation of their medical experience; (2) emotional complaints relating to the service received; and (3) suggestions for health care delivery improvement. A sample of negative comments is provided in Figure 3, whereas Table 3 provides the sample in English, accordingly. Narratives describe the medical experience honestly, whereas emotional words are highlighted in the emotion type. For suggestions, the commenter suggests possible solutions for health care delivery improvements. These three types may not always be identifiable in each comment, whereas sometimes, commenters could mention all of them. Two research assistants (RAs), both with a medical informatics background, were involved in the coding process, following a training session. They independently coded a random selection of 9.99% (301/3012) of the total comments within the pilot framework. If concepts were beyond the previous coding scheme, the two RAs discussed adjustment until consensus was reached. Any discrepancy or disagreement was solved by WZ, the lead author of this study. After the independent coding of the 301 comments, the final coding framework was formed, and the intercoder reliability was shown as Cronbach alpha=.82, indicating high credibility. Finally, one RA coded the remaining comments using the newly developed framework; the coding framework is presented in Table 4.

Figure 2. Workflow of medical consultation in a typical tertiary referral hospital.



Coding Framework

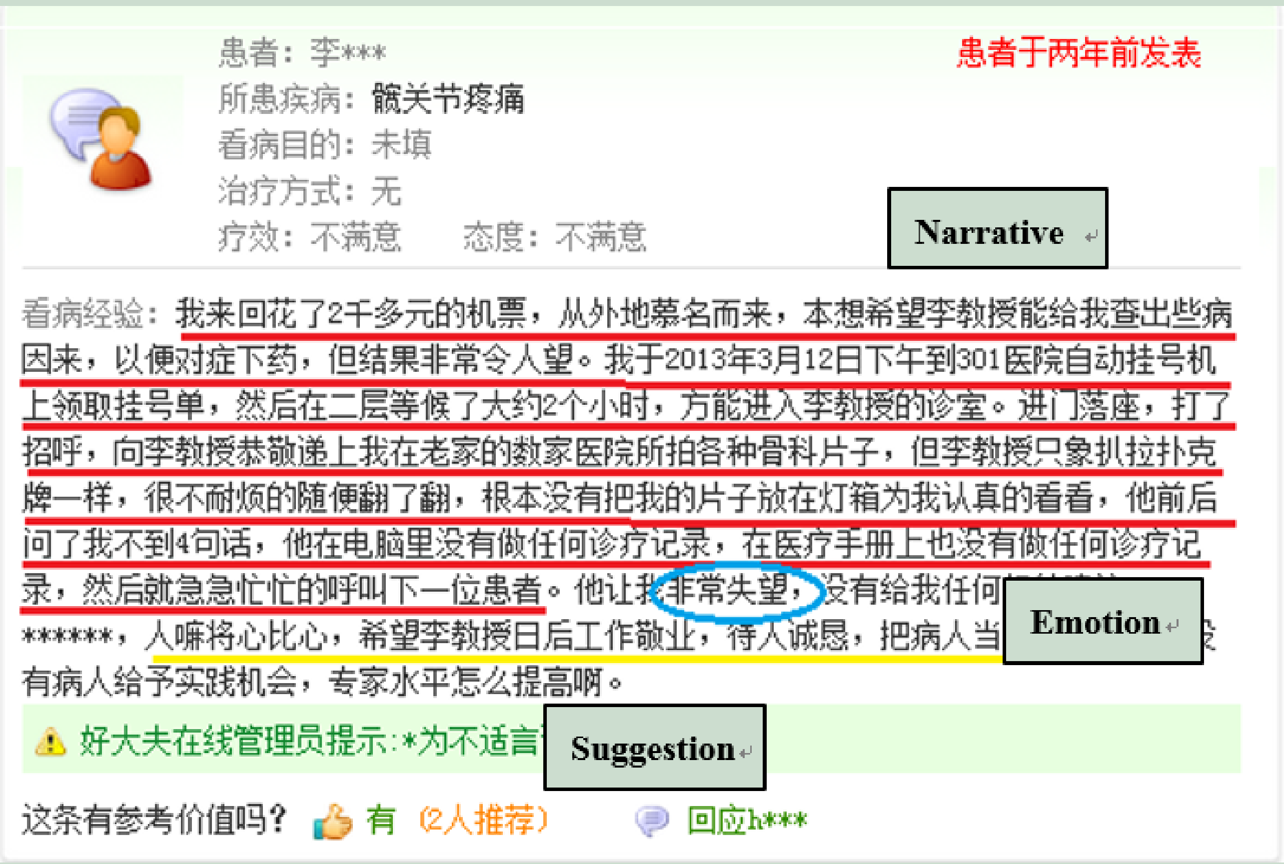
Identity

The identity of the commenters is defined by the relationship between the commenter and patients. The identity of commenters was coded into four types according to their closeness to the patient: (1) commenters as patients themselves; (2) commenters as a patient's close relative, including spouse, parent, or a grown-up child; (3) commenters as a patient's other relative; and (4) for friends and other relationships not specified. Meanwhile, some comment providers also identified themselves as "traveled patients," visiting hospitals in Beijing because of their reputation in specific medical fields. Therefore, we coded (1) for traveled patients and (0) for local patients.

Dissatisfaction

For the coding scheme relating to dissatisfaction, we followed the workflow of a health service delivery provider in a Chinese tertiary referral hospital. Generally, if a patient wants to see a doctor, he or she must register for that doctor's surgery (doctor's office) before a face-to-face consultation can occur. We coded this as the "premedical consultation" stage. During this stage, complaints include topics such as registration and waiting room issues and time taken before consultation. During medical consultation, four substages are identified, ranging from (1) overall perception, (2) preliminary diagnosis, (3) examination, to (4) the closure of consultation. Overall perception refers to the immediate evaluation of a physician's attitude and their communication with the patient. Preliminary diagnosis relates to the first contact experienced between the physician and patient, where the physician commonly employs four techniques of diagnosis, that is, to look, listen, question, and feel the pulse of the patient, to obtain an individual's information. In the examination stage, a medical device is applied to the patient, such as computed tomography film or a blood pressure monitor. The closure of consultation refers to when the patient is about to leave the hospital, with complaints typically concerning bills. After the consultation, patients may start to evaluate the effect of the treatment, if any. We coded this as "post consultation," with a focus on the patient's perception of effect. Each of the complaint areas is coded with 4 to 9 items.

Figure 3. A sample of negative comments from the Good Doctor website.



患者：李***
患者于两年前发表
所患疾病：髋关节疼痛
看病目的：未填
治疗方式：无
疗效：不满意 态度：不满意

看病经验：我来回花了2千多元的机票，从外地慕名而来，本想希望李教授能给我查出些病因来，以便对症下药，但结果非常令人望。我于2013年3月12日下午到301医院自动挂号机上领取挂号单，然后在二层等候了大约2个小时，方能进入李教授的诊室。进门落座，打了招呼，向李教授恭敬递上我在老家的数家医院所拍各种骨科片子，但李教授只象扒拉扑克牌一样，很不耐烦的随便翻了翻，根本没有把我的片子放在灯箱为我认真的看看，他前后问了我不到4句话，他在电脑里没有做任何诊疗记录，在医疗手册上也没有做任何诊疗记录，然后就急急忙忙的呼叫下一位患者。他让我非常失望，没有给我任何 ******，人嘛将心比心，希望李教授日后工作敬业，待人诚恳，把病人当 有病人给予实践机会，专家水平怎么提高啊。

⚠ 好大夫在线管理员提示:*为不适言

这条有参考价值吗？ 有 (2人推荐) 回应h***

Table 3. Three types of negative comments.

Types	Definition	Example
Narrative	Commenter describes their medical experience in an objective manner.	*I waited 2 hours to see Dr Li, but he just flipped through my results like he was playing poker for a few seconds and simply called the next patient after me in less than four sentences with me. He didn't amend any of my medical records.*
Emotion	Commenter uses emotional words to express their dissatisfaction.	*I was very disappointed and he failed to give me any good advice.*
Suggestions	Commenter provides suggestions in the hope of health care delivery improvement.	*Dr Li should learn to feel for others and work harder in the future. I hope he could be more kind to patients next time.*

Table 4. Coding framework for negative comments.

Stages	Complaint areas	Code
Premedical consultation	Registration	1=hard to register, 2=long waiting hours, 3=high cost, 4=others
	Waiting for consultation	1=long waiting, 2= chaotic queuing
Medical consultation	**Overall perception**	
	Physician’s attitude	1=impatience, 2=disrespect patients, 3=not caring patients, 4= unavailability of physicians on duty, 5=do irrelevant things, 6=others
	Patient-physician communication	1=lacking communication, 2=not answering, 3=no time for communication, 4=others
	Preliminary diagnosis	1=ignorance of medical records and previous reports, 2=no observation (dermatology issues), 3=no inquiries, 4=others
	Examinations	1=lacking basic examinations, 2=too many examinations, 3=rude examinations, 4=repeated and inappropriate examinations, 5=long wait hours for the results, 6=no analysis for the results, 7=high cost. 8=privacy issue, 9=others
	Closure of consultation	1=no lifestyle advice, 2=no analysis before medical advice, 3=no diagnosis conclusion, 4=high cost of medical, 5=short time for diagnosis, 6=no treatment plan, 7=misdiagnosis
Post consultation	Patient’s perception of effect	1=no effect or little effect, 2=worse than before, 3=inappropriate treatment plan, 4=others

Results

The Majority of Commenters Were Patients Themselves, With Several Comments Being Contributed by Accompanying Persons During Medical Consultation

A total of 86.22% (2597/3012) of commenters shared their experience as patients, contributing direct feedback relating to the perceived effectiveness and efficiency of the medical service they received (see Figure 4 for details). For the accompanying persons, the largest group were the patient’s grown-up children (227/3012, 7.54%), followed by the patient’s parents (149/3012, 4.95%). With the aging population of China, the country has a high percentage of “older” inhabitants, with the chance of this age group contracting diseases, particularly chronic illnesses, being considerably higher than their younger counterparts. It is common for young adults in China to accompany their older parents when visiting a doctor. Meanwhile, taking care of a child’s health is also an important responsibility in Chinese family culture, and parents typically accompany their children during hospital visits and treatment.

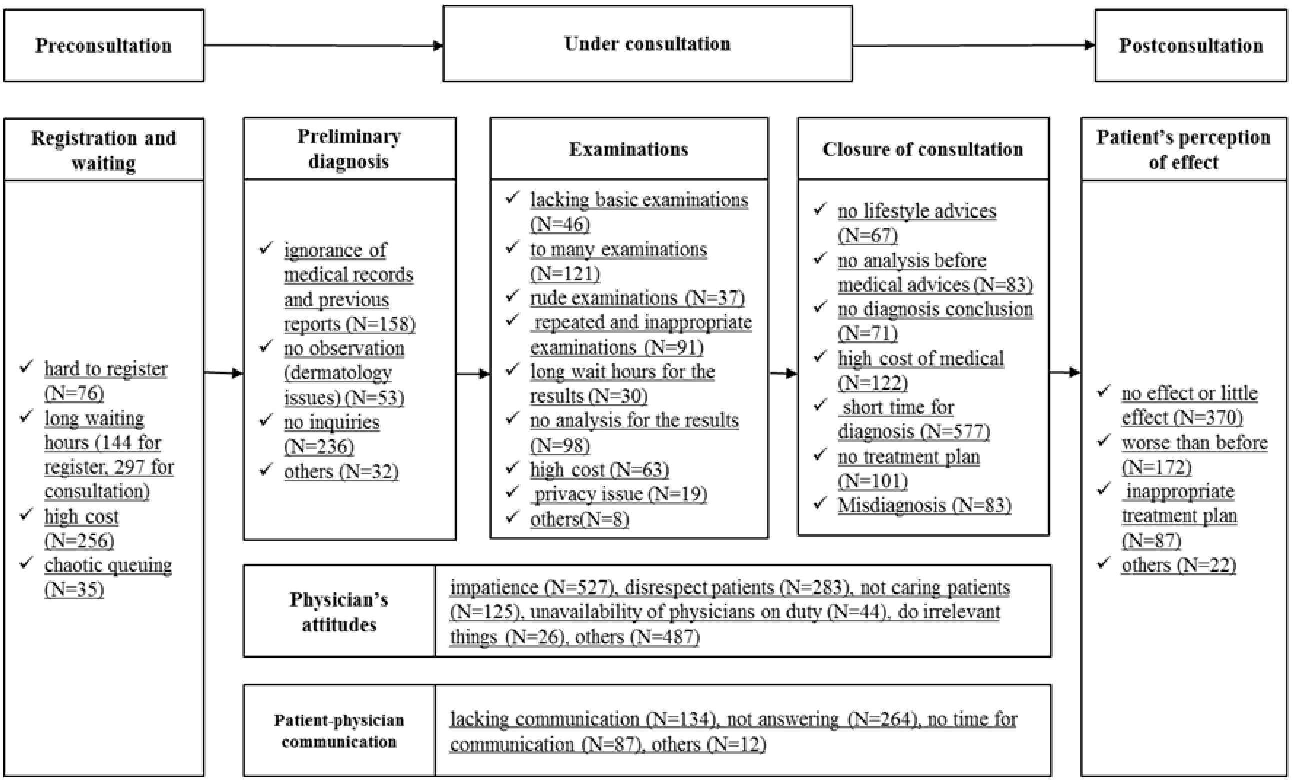
In addition, 228 (7.57%, 228/3012) dissatisfied commenters were “traveled patients.” This group had traveled to one of the five highly rated hospitals in Beijing because of the excellent level of medical services perceived to be provided there. These patients are more concerned with the quality of medical consultation received because of the additional costs incurred from the long distance traveled and accommodation required.

Many Complaints Posted Related to Long Waiting Times Experienced and High Costs Associated to Registration During the Premedical Consultation Stage

Of all negative comments mentioned during the premedical consultation stage, long waiting times accounted for the majority of complaints, representing 7.9% (350/441) of total issues. It is common for patients to feel upset and/or anxious when suffering from a poor health condition that would cause them to have little tolerance for a long waiting time. High costs relating to registration fees is second, whereas 29 comments highlighted a special group—ticket touts. Ticket touts, or *Huangniu* in Chinese, in the context of hospital settings, refer to an individual or group of people who work as agents for patients to acquire their desired, but usually hard to obtain, reservation tickets to see renowned doctors. If ticket touts are able to obtain the required tickets, they may charge the patients much more than the normal ticket price. For example, a reservation ticket costs between 7 to 14 yuan but could sell for at least several hundred yuan and sometimes thousands. On the Good Doctor website, patients expressed their dissatisfaction with this unethical practice. They believed that ticket touts have disturbed the regular pricing strategy and cause increased difficulties during registration. Hard to register (N=76) was also specified. Cutting in lines or disorderly queuing systems (N=35) may further worsen the situation. One grown-up child accompanying their parent noted the following:

> *We have tried quite a lot to reserve Doctor X. We waited for a whole night to get registered...I do not know what happened to the auto-calling queuing system and we waited until 11 in the morning though there were still quite a few patients ahead of us.*

Figure 4. Main results of the coding.



Physician Attitudes Toward Patients Are of Great Concern, With Impatience and Disrespect Being Identified Most Often

The attitude of physicians toward patients in medical settings can be considered crucial, with a total of 1492 complaints (26.91%, 1492/5544) identified. Among them, the impatience of physicians was mentioned most frequently (527/5544, 9.51%), with impatience relating to physician responses to patient questions, explanations of medical results and daily communication with patients being identified. Impatience was perceived by commenters as emotional abuse and could exert substantial influence on a patient's mood. Although the comments mainly concerned the attitude of physicians, 17 commenters also blamed the physician's assistants or medical staff for poor attitude. The disrespecting of patients came second (283/5544, 5.10%); disrespectful behaviors are distinguished from general bad attitudes, which encompass detesting, criticizing, and/or blaming of patients, with or without verbal abuse. Some "traveled patients" complained that their physicians demonstrated regional discrimination and perceived unfair treatment compared with local inhabitants of Beijing. In addition, another 125 (4.15%, 125/3012) comments emphasized little care perceived from physicians. Patients can be considered emotionally fragile and sensitive and may need more attention than what has previously been experienced. One patient commented the following:

> *We went to see Doctor Sun in the anus-intestines Department...he just asked us a few questions with a bad attitude. We felt awful because of his impatience...We hope that Dr Sun could think in the patients' shoes and have a parental heart.*

Poor Physician-Patient Communication Was Highlighted by Patients and Their Accompanying Persons

Due to the knowledge gap between physicians and patients, physicians typically dominate the conversations and may not involve patients in decision making. A total of 264 (8.76%, 264/3012) commenters posted negative comments based on the physician not answering their questions, whereas 87 commenters felt that they experienced no communication between themselves and the physician. In addition, another 134 (4.45%, 134/3012) posted complaints about insufficient time spent communicating. One commenter wrote about their regret when their mother was diagnosed with synovitis, as follows:

> *Five months after the diagnosis...we waited for a long time and managed to see the doctor around 5 PM. We wanted to talk with the doctor, but it seemed that he did not want to talk with us at all. He did not want to answer our questions and demanded us to pack the film. He interrupted our several attempts to ask questions and we had to give up...We felt helpless as a patient.*

No Inquiry From Physicians Caused Serious Dissatisfaction During Preliminary Diagnosis

The preliminary diagnosis stage is the first form of contact experienced between the patient and physician before a further examination with medical devices is conducted. At this stage, physicians usually perform a series of diagnostic tasks such as looking, smelling, questioning, and taking the pulse of the patient to find out the cause, severity, or development of the disease or symptoms. A total of 236 (7.84%, 236/3012) commenters posted that their physician had not inquired about their disease or symptoms, whereas another 158 (5.25%, 158/3012) mentioned that no concern for medical records or previous reports was shown. These behaviors could lead to an unprofessional image of the health care providers. Patients may also feel that physicians are too busy to draw an informed conclusion without collecting physical evidence, which may lead to a reluctance to follow physician advice. One patient noted the following:

> *My son has had neck ache for years. I came for his reputation in pediatrics. However, he just glimpsed at the CT film for a second and did not inquire about my kid's condition at all...I felt that he never cared and just concluded that my kid was OK. That's really unacceptable.*

Too Many Examinations Conducted and No Analysis of Results Were Major Factors Associated With Dissatisfaction During the Examination Stage

A total of 121 (4.02%, 121/3012) commenters stated that the number of examinations conducted on them were too many, whereas 91 patients thought that their examinations were repeats of previous examinations or unnecessary. Increased examinations mean high costs for patients, and 63 commenters believed that the chaotic arrangement of examinations resulted in them paying more money. In addition, 98 commenters expressed unhappiness because of their physicians not analyzing their results. A total of 30 commenters felt that the waiting time for their reports was unacceptable. Regarding privacy issues experienced during examinations, gynecology patients constantly mentioned the problem of more than one patient being jointly examined at the same time. Rude behavior such as too much strength or pressure used during examination was also reported in 37 cases. One commenter noted the following:

> *I went to the doctor last year. He asked me to go through a series of checks. He instructed me to take CT films and B-ultrasound each, twice. Others, like gastroscopy...the staff examined me as if I were an object...It was a hurtful experience in that hospital.*

Limited Time Spent in Medical Consultation Was the Main Complaint Reported on at the End of the Medical Service

A total of 577 (19.16%, 577/3012) commenters discussed the duration of medical consultation, which may explain the perceived "rushed" attitude of physicians and the insufficient communication experienced. Further exploration reveals that patients are more likely to post negative comments if long waiting times and short medical consultation time are experienced together. One commenter noted the following:

> *My father has insomnia...and I decided to take him to see this doctor. We have waited for 3 hours, but the whole medical session did not last longer than 8 minutes. I felt so upset.*

Furthermore, 122 (4.05%, 122/3012) patients complained about costs associated to medical fees. Another 101 (3.35%, 101/3012) commenters were disappointed because of no advice being offered during treatment, whereas 83 patients found it hard to accept their medical schedule without any medical analysis being explained to them. A total of 83 patients were dissatisfied because of not receiving a conclusive diagnosis by the end of their consultation.

Therapeutic Effect Was Considered the Most Important Experience During Postmedical Consultation

Following consultation, patients cared most about the improvement in their health condition. A total of 370 (12.28%, 370/3012) patients rated therapeutic effect as unsatisfactory because of the results they expected post consultation. A total of 172 (5.71%, 172/3012) commenters mentioned that their health condition had become worse since their premedical consultation. In addition, 87 patients believed that the medical schedule advised may be inappropriate. Some claimed that after consulting another physician, who made a different diagnosis, their disease was eventually cured. It is highly possible that patients feel unstable and give a low rating to physicians if they perceive little therapeutic effect. Sometimes, therapeutic effect may outweigh attitude problems experienced. For example, one patient mentioned the following:

> *Dr Song is very careful and has a perfect attitude toward us, but he cannot solve our problem. A simple allergic rhinitis, he fails to diagnose, let alone the other ones. He just gave me some vitamins and sent me home.*

Discussion

Dissatisfied or Unhappy Patients Are Not Always Alike, With Reasons for Feeling Dissatisfied Being Individual to Each Patient

Patient dissatisfaction is experienced throughout all stages of medical consultation. Although efforts have been made to classify all negative factors that appeared in the 3012 comments, it is still extremely difficult to present them with limited codes. Many complaints are intertwined; for example, perceived high costs can be associated with either registration, at the examination stage, or as part of the closure of consultation. If patients perceive that the medical consultation was expensive, they are more likely to complain about the short duration of consultation, the physician's bad attitude, and/or poor therapeutic effect. However, for reputable tertiary referral hospitals in Beijing, a comparatively higher cost should be expected compared with ordinary hospitals; this is because of patients having higher expectations when they visit physicians in reputed hospitals, as they hope doctors can solve their problems that have not been addressed at other less reputable hospitals. With high expectations, several groups of patients are seen to be difficult to satisfy, including patients with chronic or complicated diseases, "traveled patients," and very important person registrants. Meanwhile, accompanying persons who have strong feelings for their beloved ones, usually old parents or children, are easily irritated [28]. For both physicians and the hospitals, special attention is required for these groups of patients to avoid possible physician-patient conflicts [29].

Interestingly, we found that the Departments of Obstetrics and Gynecology and Internal Medicine received the most negative comments. This may be because of physicians in these departments receiving the highest number of comments in general on Good Doctor website [14,19]. In addition, the department staff also imply that their patients may be hard to please. For the former, perinatal care usually involves more family members, and their expectancy of a new baby makes them more sensitive to the medical care provided, especially for the primipara. For the latter, as Internal Medicine is generally difficult to observe, compared with Surgeries, patients may be more critical and easily irritated; this could also explain why departments such as Dermatology and STD and Andrology were the most complained about department per physician. Specifically, dermatology and plastic surgery correlate with each other, and patients may require an increased time to become accustomed to physical and mental changes; for STD or Andrology, no immediate cures are available, and patients have a strong concern about their privacy.

Patient Complaints Are Complicated and Contradictory

With the diversity of patient complaints recorded, it is evident that patient perceptions are often different. For example, for the same physician, one patient may perceive none or few examinations as a “rushed” service to make decisions and move on to the next patient, whereas others may loathe excessive or inappropriate examinations. An example of this is *Jiahao* (adding reexamination patients), where physicians will add patients for reexamination to the top of their patient list, saving time for the patient. However, some patients regard *Jiahao* as unfair and perceive that it could result in much longer waiting times than usual. Since the duration of medical consultation for a physician is a fixed period every day, if the physician sees one more patient, it means less time for another patient. In this regard, it brings us to the issue of “justice,” in the context of medical services. Each patient may wish greater attention is given to them by their physician, so it is often difficult for them to tolerate other patients who take up a physician’s time. If this happens, the patient may perceive that they have been unfairly treated and blame the physician because of them having the power to manage the time of each appointment. However, it is unrealistic to implement a rule that the physician must spend a certain amount of time with a patient because of patient diversity, such as severity and types of diseases. In general, each patient’s preferences and understanding of a medical consultation is different, and no simple explanation for each patient’s dissatisfaction can be identified. Furthermore, although patients tend to avoid physicians with negative comments, it is sometimes unavoidable because of various reasons. One comment revealed the following:

> *I searched almost all the doctors who might solve my problems online and found everyone said X was not a good one. I wish I could avoid her, however the only one I could see is her upon my arrival.*

Incorrect Medical Advice, Overcrowded Medical Resources, and Policy Failures in Beijing Hospitals Jointly Contribute to the Dissatisfaction of Patients

For most Chinese people, modern medical treatment is perceived to be lifesaving, using the technologies or skills possessed by physicians. Many Chinese people believe that hospitals and their physicians should solve all patient problems; if not, they perceive the hospital or physician is not good enough. With this in mind, patients believe that the best, most highly rated hospitals or physicians are rare, and their health conditions may be best addressed, as long as they are willing to pay more. In the case of tertiary referral hospitals in Beijing, people are convinced that they are the best hospitals in China and wish that they can go to the best hospitals whenever they are sick. As a result, patients expect hospitals to be overcrowded, with increased waiting times and physicians spending less time with each patient. However, modern medical treatment does not provide a cure-all approach. A patient with the common flu will take approximately 1 week to recover. Seeing the best doctor in the best hospital will not reduce this time and simply deprives those who require medical assistance most of the medical resources. In recent years, China has started to implement a hierarchical medical scheme; this aims to divert patients in tertiary referral hospitals to primary hospitals and community hospitals. However, this policy is still in its infancy. Most people distrust physicians in primary or community hospitals and feel reluctant to see a doctor in these hospitals [30]; this causes an unexpected influx of patients at tertiary referral hospitals, which must satisfy emerging health demands from the public. With the increasing prosperity of China, Chinese people are now more concerned about their health conditions and have more demands or expectancy toward physicians that is often very difficult to satisfy.

Blame the Physicians, but Also Try to Sympathize With Them

Throughout the medical consultation, physician-patient communication is still a key factor that affects patient satisfaction. Patients are human beings, and their perception of medical consultation is mainly built upon the communication had with their physicians. For the physicians, a warm heart and a good, friendly attitude is very necessary [31]. In China, hostility between physicians and patients is largely caused by the limited medical resources available to the physician [32]. These limited resources cannot satisfy the emerging demands for health care, with medical agents such as ticket touts causing extra strain on the delivery of medical services. In the case of ticket touts, they usually sell their “tickets” through one of four approaches: (1) offline queuing. Ticket touts are familiar with the rules of ticket delivery and arrive early to ensure the first position in the ticket queue. Their appointment ticket can then be sold to any paying patient, (2) tickets from internal resources. Ticket touts take full advantage of their *guanxi* or good relationship with physicians, nurses, and internal staff to obtain internal or additional tickets. It is sometimes possible that internal staff may collaborate with ticket touts, (3) ticket-buying plug-ins or mobile apps. As an increasing number of hospitals have embraced the Internet for ticket delivery, ticket touts provide paid services for making appointments for patients online, and (4) stocking up and reselling tickets. For this approach, the ticket touts need to find bugs or “loopholes” in the Web-based ticket-buying system. For example, suppose one physician’s tickets for today were sold out online, the ticket tout could promise an additional ticket as they have reserved extra tickets. If they cancel some of them, new spaces will appear for the patients within several minutes. On the patients’ side, they perceive there to be a high possibility that physicians collaborate with the ticket touts. In extraordinary cases, people may dress up in medical gowns in large hospitals and pretend that they are reputable physicians, soliciting money from those who desperately need health care services in exchange for fictitious advice. These experiences lead to a common distrust toward physicians.

In one sense, physicians must be responsible for patients and improve their quality of health care. From another perspective, patients are expected to understand their physicians [33]. We inferred from some patient comments that the impatience of physicians may contribute to the overwhelming number of patients and called for mutual understanding between patients and physicians. Though Web-based negative comments will not disappear immediately [34], physicians may not worry too much about them as many patients make comments on impulse when feeling frustrated by their experience [35]. Instead, complaints could be viewed as free advice for both the hospital and physician to enhance the quality of health care provision [36,37]. Thanks to the anonymity and convenience of expressing dissatisfaction online, patients can evaluate their physicians more precisely without too much consideration being given to social context, such as obeying complex social and cultural norms [36]. From this perspective, physicians could be more open to negative comments and learn from their own failures in health care delivery to make further improvement.

Limitations

This study is not without flaws. First, data were only captured from the locality of Beijing on the Good Doctor website. It is possible that this data only reflects the dissatisfaction for specific online users, and the conclusions may not apply to other small- and medium-sized cities or hospitals. Meanwhile, as medical rating websites have their own bespoke functionality, comments collected may suggest different outcomes. Second, because of the complexity of the Chinese language, where words have dual meanings, our analysis framework may omit some potential attitudes and complaints. Third, it is assumed that negative comments relate to a patient’s real experience. However, it is possible that negative comments are manipulated by competitors. Future research could consider a further qualitative approach; for example, focus group or in-depth interview methods could be used with patients who have rated their physician negatively online. Additionally, the mechanism of how online negative comments affect patient decision making requires more attention from researchers [28,38].

Conclusions

This study is different from those that have focused on all Web-based comments for hospitals or physicians [39]; in contrast, we strived to explore the factors associated with patient dissatisfaction through rigorous content analysis of negative patient comments.

Due to increasing medical specialization, patients are in a comparatively disadvantaged position compared with physicians. Research on patient vulnerability through factors associated with dissatisfaction is crucial to the quality in delivery of health care services and patient safety. Despite the prevalence of on-site complaints in medical institutions, few patients adopt this approach to voice their complaints. To analyze the factors associated with dissatisfaction, we collected self-reported patient experiences toward certain physicians on the Chinese medical platform, the Good Doctor website. Though comments are expected to be centered on the physician, patients also discussed their overall experience, covering a wide range of issues, including hospital registration and the attitude of nurses or other staff members, to help other patients choose the right physician, hospital, and medical treatment.

Finally, we conducted content analysis to explore the negative comments of patients and found that all patients have individual and unique concerns. Negative factors are identified in almost all stages of the medical consultation. The factors are often connected but also distinct from one other. In addition, individual variations make these factors more complex. Among them, the key complaints received are limited medical consultation time and impatience of physicians. Other complaints include the level of therapeutic effect experienced, poor treatment schemes, incorrect information being provided, and disrespect toward patients. This is somewhat consistent with previous studies on the analysis of Web-based and offline complaints toward physicians [18,40], but with more detail concerning the workflow of health care delivery. Meanwhile, it should be noted that physician behavior is shaped greatly by the national health care allocation and health care system [41]; no simple solution for the improvement of patient satisfaction or sustainable behavior adjustment exists—this requires fundamental change in the Chinese health care system.

Acknowledgments

This work has been partially supported by the National Natural Science Foundation of China (Award #71671073), Hubei Social Science Foundation (Award #2017188), and the Fundamental Research Funds for the Central Universities (Award #2015AB019).

Conflicts of Interest

None declared.

References

1. Hsieh SY. A system for using patient complaints as a trigger to improve quality. Qual Manag in Health Care 2011;20(4):343-355. [doi: 10.1097/QMH.0b013e318222e73b] [Medline: 21971031]
2. Howard M, Fleming ML, Parker E. Patients do not always complain when they are dissatisfied: implications for service quality and patient safety. J Patient Saf 2013;9(4):224-231. [doi: 10.1097/PTS.0b013e3182913837] [Medline: 24257066]
3. Phau I, Baird M. Complainers versus non-complainers retaliatory responses towards service dissatisfactions. Mark Intell Plan 2008;26(6):587-604. [doi: 10.1108/02634500810902848]
4. Gurung G, Derrett S, Gauld R, Hill PC. Why service users do not complain or have 'voice': a mixed-methods study from Nepal's rural primary health care system. BMC Health Serv Res 2017 Jan 25;17(1):81 [FREE Full text] [doi: 10.1186/s12913-017-2034-5] [Medline: 28122552]
5. Jiang Y, Ying X, Zhang Q, Tang SR, Kane S, Mukhopadhyay M, HESVIC team authorship. Managing patient complaints in China: a qualitative study in Shanghai. Br Med J Open 2014;4(8):e005131. [doi: 10.1136/bmjopen-2014-005131]
6. Skålén C, Nordgren L, Annerbäck EM. Patient complaints about health care in a Swedish County: characteristics and satisfaction after handling. Nurs Open 2016;3(4):203-211. [doi: 10.1002/nop2.54]
7. Sparks BA, Browning V. Complaining in cyberspace: the motives and forms of hotel guests' complaints online. J Hosp Mark Manage 2010;19(7):797-818. [doi: 10.1080/19368623.2010.508010]
8. Patel S, Cain R, Neailey K, Hooberman L. Exploring patients' views toward giving web-based feedback and ratings to general practitioners in England: a qualitative descriptive study. J Med Internet Res 2016;18(8):e217. [doi: 10.2196/jmir.5865]
9. 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]
10. Reimann S, Strech D. The representation of patient experience and satisfaction in physician rating sites. A criteria-based analysis of English- and German-language sites. BMC Health Serv Res 2010;10:332 [FREE Full text] [doi: 10.1186/1472-6963-10-332] [Medline: 21138579]
11. 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 Nov 16;13(4):e95 [FREE Full text] [doi: 10.2196/jmir.1960] [Medline: 22088924]
12. López A, Detz A, Ratanawongsa N, Sarkar U. What patients say about their doctors online: a qualitative content analysis. J Gen Intern Med 2012;27(6):685-692. [Medline: 22215270]
13. Emmert M, Meier F, Heider AK, 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;118(1):66-73. [doi: 10.1016/j.healthpol.2014.04.015] [Medline: 24836021]
14. Hao H. The development of online doctor reviews in China: an analysis of the largest online doctor review website in China. J Med Internet Res 2015;17(6):e134. [Medline: 26032933]
15. Brody S, Elhadad N. Detecting salient aspects in online reviews of health providers. AMIA Annu Symp Proc 2010;2010:202-206.
16. 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]
17. Rastegar-Mojarad M, Ye Z, Wall D, Murali N, Lin S. Collecting and analyzing patient experiences of health care from social media. JMIR Res Protoc 2015 Jul 02;4(3):e78 [FREE Full text] [doi: 10.2196/resprot.3433] [Medline: 26137885]
18. Doing-Harris K, Mowery DL, Daniels C, Chapman WW, Conway M. Understanding patient satisfaction with received healthcare services: a natural language processing approach. AMIA Annu Symp Proc 2017;2016:524-533 [FREE Full text] [Medline: 28269848]
19. Hao H, Zhang K. The voice of Chinese health consumers: a text mining approach to web-based physician reviews. J Med Internet Res 2016;18(5):e108. [Medline: 27165558]
20. Wallace BC, Paul MJ, Sarkar U, Trikalinos TA, Dredze M. A large-scale quantitative analysis of latent factors and sentiment in online doctor reviews. J Am Med Inform Assoc 2014;21(6):1098-1103 [FREE Full text] [doi: 10.1136/amiajnl-2014-002711] [Medline: 24918109]
21. Li S, Feng B, Chen M, Bell RA. Physician review websites: effects of the proportion and position of negative reviews on readers' willingness to choose the doctor. J Health Commun 2015;20(4):453-461. [Medline: 25749406]
22. Gordon HS. Consider embracing the reviews from physician rating websites. J Gen Intern Med 2017;32(6):599-600. [doi: 10.1007/s11606-017-4036-8] [Medline: 28299603]
23. Patel S, Cain R, Neailey K, Hooberman L. General practitioners' concerns about online patient feedback: findings from a descriptive exploratory qualitative study in England. J Med Internet Res 2015 Dec 08;17(12):e276 [FREE Full text] [doi: 10.2196/jmir.4989] [Medline: 26681299]
24. Haodf. The Good Doctor URL: http://www.haodf.com/ [accessed 2018-01-17] [WebCite Cache ID 6wXOJpn5d]
25. Zhang W, Xu X, Zhang H, Chen Q. Online participation chaos: a case study of Chinese government-initiated e-polity square. Int J Public Adm 2016;39(14):1195-1202. [doi: 10.1080/01900692.2015.1072218]
26. Montini T, Noble AA, Stelfox HT. Content analysis of patient complaints. Int J Qual Health Care 2008;20(6):412-420. [doi: 10.1093/intqhc/mzn041] [Medline: 18801751]

27. Reader TW, Gillespie A, Roberts J. Patient complaints in healthcare systems: a systematic review and coding taxonomy. Br Med J Qual Saf 2014;23(8):678-689 [FREE Full text]
28. Burt J, Abel G, Elmore N, Lloyd C, Benson J, Sarson L, et al. Understanding negative feedback from South Asian patients: an experimental vignette study. Br Med J Open 2016 Dec 08;6(9):e011256 [FREE Full text] [doi: 10.1136/bmjopen-2016-011256] [Medline: 27609844]
29. [No authors listed]. Violence against doctors: Why China? Why now? What next? Lancet 2014 Mar 22;383(9922):1013. [doi: 10.1016/S0140-6736(14)60501-8] [Medline: 24656183]
30. Tucker JD, Cheng Y, Wong B, Gong N, Nie JB, Zhu W, Patient-Physician Trust Project Team. Patient-physician mistrust and violence against physicians in Guangdong Province, China: a qualitative study. Br Med J Open 2015 Oct 06;5(10):e008221 [FREE Full text] [doi: 10.1136/bmjopen-2015-008221] [Medline: 26443652]
31. Stefani I, Teti O, Mazzone A. Person-centred analysis in physician-patient relationship. Eur J Intern Med 2017(43):e26-e27. [doi: 10.1016/j.ejim.2017.05.014] [Medline: 28549822]
32. Chan CS, Cheng Y, Cong Y, Du Z, Hu S, Kerrigan A, et al. Patient–physician trust in China: a workshop summary. The Lancet 2016 Oct;388:S72. [doi: 10.1016/S0140-6736(16)31999-7]
33. Alders I, Smits C, Brand P, van Dulmen S. Does patient coaching make a difference in patient-physician communication during specialist consultations? A systematic review. Patient Educ Couns 2017;100(5):882-896. [doi: 10.1016/j.pec.2016.12.029] [Medline: 28089309]
34. Pasternak A, Scherger JE. Online reviews of physicians: what are your patients posting about you? Fam Pract Manag 2009;16(3):9-11 [FREE Full text] [Medline: 19492765]
35. Gao GG, McCullough JS, Agarwal R, Jha AK. A changing landscape of physician quality reporting: analysis of patients' online ratings of their physicians over a 5-year period. J Med Internet Res 2012 Feb 24;14(1):e38 [FREE Full text] [doi: 10.2196/jmir.2003] [Medline: 22366336]
36. Bomhoff M, Friele R. Complaints in long-term care facilities for older persons: why residents do not give 'free advice'. Health Policy 2017;121(1):75-81. [doi: 10.1016/j.healthpol.2016.11.007] [Medline: 27889081]
37. Hawkins JB, Brownstein JS, Tuli G, Runels T, Broecker K, Nsoesie EO, et al. Measuring patient-perceived quality of care in US hospitals using Twitter. Br Med J Qual Saf 2016 Jun;25(6):404-413 [FREE Full text] [doi: 10.1136/bmjqs-2015-004309] [Medline: 26464518]
38. Grabner-Kräuter S, Waiguny MK. Insights into the impact of online physician reviews on patients' decision making: randomized experiment. J Med Internet Res 2015 Apr 09;17(4):e93 [FREE Full text] [doi: 10.2196/jmir.3991] [Medline: 25862516]
39. Lagu T, Goff SL, Hannon NS, Shatz A, Lindenauer PK. A mixed-methods analysis of patient reviews of hospital care in England: implications for public reporting of health care quality data in the United States. Jt Comm J Qual Patient Saf 2013 Jan;39(1):7-15. [Medline: 23367647]
40. Schaad B, Bourquin C, Bornet F, Currat T, Saraga M, Panese F, et al. Dissatisfaction of hospital patients, their relatives, and friends: analysis of accounts collected in a complaints center. Patient Educ Couns 2015 Jun;98(6):771-776. [doi: 10.1016/j.pec.2015.02.019] [Medline: 25800651]
41. Tang S, Brixi H, Bekedam H. Advancing universal coverage of healthcare in China: translating political will into policy and practice. Int J Health Plan Manage 2014;29(2):160-174. [doi: 10.1002/hpm.2207] [Medline: 23877959]

Abbreviations

GP: general practitioner
PRW: physician rating websites
STD: sexually transmitted diseases

Edited by N Collier, M Conway, N Limsopatham, V Lampos, A Culotta; submitted 16.06.17; peer-reviewed by S Pavlinovic, M Conway; comments to author 01.09.17; revised version received 16.10.17; accepted 01.12.17; published 25.01.18.

Please cite as:
Zhang W, Deng Z, Hong Z, Evans R, Ma J, Zhang H
Unhappy Patients Are Not Alike: Content Analysis of the Negative Comments from China's Good Doctor Website
J Med Internet Res 2018;20(1):e35
URL: http://www.jmir.org/2018/1/e35/
doi:10.2196/jmir.8223
PMID:29371176

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org